OM nucleic - nucleic search, using sw model

          GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: March 22, 2001, 01:35:40 ; Search time 1698.78 Seconds
                                           (without alignments)
                                           4202.582 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaataccaata........tcaggcaagtccccgtcaa 1395

Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0

Searched: 118133 seqs, 2558875100 residues

Word size :   25

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   GenEmbl:*
          1: gb_ba1:*
          2: gb_ba2:*
          3: gb_om:*
          4: gb_ov:*
          5: gb_ph:*
          6: gb_pl1:*
          7: gb_pl2:*
          8: gb_pr1:*
          9: gb_pr2:*
         10: gb_pr3:*
         11: gb_ro:*
         12: gb_sy:*
         13: gb_un:*
         14: em_fun:*
         15: em_hum1:*
         16: em_hum2:*
         17: em_in:*
         18: em_om:*
         19: em_or:*
         20: em_ov:*
         21: em_pat:*
         22: em_ph:*
         23: em_pl:*
         24: em_ro:*
         25: em_sts:*
         26: em_sy:*
         27: em_un:*
         28: em_vi:*
         29: gb_ba3:*
         30: gb_in1:*
         31: gb_in2:*
         32: gb_in3:*
         33: gb_pl3:*
         34: gb_pr4:*
         35: gb_ba1:*
         36: gb_ba2:*
         37: em_htg1:*
         38: em_htg2:*
         39: em_htg3:*
         40: em_htg4:*
         41: em_htg5:*
         42: em_htg6:*
         43: em_htg7:*
         44: em_htg8:*
         45: em_htg9:*
         46: em_htg10:*
         47: em_hum3:*
         48: em_hum4:*
         49: em_hum5:*
         50: em_hum6:*
         51: gb_pr5:*
         52: gb_pr6:*
         53: gb_pr7:*
         54: gb_htg1:*
         55: gb_htg2:*
         56: gb_htg3:*
         57: gb_htg4:*
         58: gb_htg5:*
         59: gb_htg6:*
         60: gb_htg7:*
         61: gb_htg8:*
         62: gb_htg9:*
         63: gb_htg10:*
         64: gb_htg11:*
         65: gb_htg12:*
         66: gb_htg13:*
         67: gb_htg14:*
         68: gb_htg15:*
         69: gb_htg16:*
         70: gb_htg17:*
         71: gb_htg18:*
         72: gb_htg19:*
         73: gb_htg20:*
         74: gb_htg21:*
         75: gb_htg22:*
         76: gb_htg23:*
         77: gb_sts1:*
         78: gb_sts2:*
         79: gb_vi1:*
         80: gb_vi2:*
         81: gb_pat1:*
         82: gb_pat2:*
         83: em_htg0:*
         84: gb_htg24:*
         85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

                    %
Result   Query
No.   Score   Match   Length   DB   ID   Description
----- -----   -----   ------   --   --------   -----------
  1   175   12.5   1500   2   AX011572   AX011572 Sequence
  2   175   12.5   1500   2   AX011574   AX011574 Sequence
c 3   175   12.5   311321   2   NMA3Z2491 AL162754 Neisseria
  4   156   11.2   1500   2   AX011576   AX011576 Sequence
c 5   156   11.2   10531   1   AE002409   AE002409 Neisseria
  6   113   8.1   1110   2   AX011578   AX011578 Sequence
  7   109   7.8   390   2   AX011580   AX011580 Sequence

                              ALIGNMENTS

RESULT   1
AX011572
LOCUS   AX011572   1500 bp   DNA   BCT   18-SEP-2000
DEFINITION   Sequence 1 from Patent WO9955872.
ACCESSION   AX011572
VERSION   AX011572.1   GI:9998105
KEYWORDS   .
SOURCE   Neisseria meningitidis.
  ORGANISM   Neisseria meningitidis

```
REFERENCE   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1500)
AUTHORS     Ruelle,J.L.
TITLE       Basb013 dna and proteins from neisseria meningitidis
JOURNAL     Patent: WO 955872-A  04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1..1500
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      370 a    468 c    393 g    269 t
ORIGIN

Query Match          12.5%;  Score 175;  DB 2;  Length 1500;
Best Local Similarity 99.6%;  Pred. No. 2.8e-84;
Matches 225;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    174 ccaactggttcaaagcggaaggcccggcagtgtcaatattcaggcagccccgccccgcg 233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    174 CCAACTGGTTCAAAGCGGAAGGCCCGGCAGTGTCAATATTCAGGCAGCCCCGCCCCGCG 233

Qy    234 cacccaaaacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgaccc 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    234 CACCCAAAACGGCAGCGGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCC 293

Qy    294 gttctacgaattttcaaacgcctcgtcccgaacatgcccgaaatgccccaagaagaagc 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    294 GTTCTACGAATTTTCAAACGCCTCGTCCCGAACATGCCCGAAATGCCCCAAGAAGAAGC 353

Qy    354 agatgacgcggcggattgaacttcggttcggggcttcatcatcagcaaa 399
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    354 AGATGACGCGGCGGATTGAACTTCGGTTCGGGGCTTCATCATCAGCAAA 399
```

```
RESULT  2
AX011574
LOCUS       AX011574      1500 bp    DNA             BCT       18-SEP-2000
DEFINITION  Sequence 3 from Patent WO955872.
ACCESSION   AX011574
VERSION     AX011574.1   GI:9998106
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1500)
AUTHORS     Ruelle,J.L.
TITLE       Basb013 dna and proteins from neisseria meningitidis
JOURNAL     Patent: WO 955872-A  04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1..1500
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      370 a    469 c    392 g    269 t
ORIGIN

Query Match          12.5%;  Score 175;  DB 2;  Length 1500;
Best Local Similarity 99.6%;  Pred. No. 2.8e-84;
Matches 225;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    174 ccaactggttcaaagcggaaggcccggcagtgtcaatattcaggcagccccgccccgcg 233
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    174 CCAACTCGTTCAAAGCGGAAGGCCCGGCAGTGTCAATATTCAGGCAGCCCCGCCCCGCG 233

Qy    234 cacccaaaacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgaccc 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    234 CACCCAAAACGGCAGCGGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCC 293

Qy    294 gttctacgaattttcaaacgcctcgtcccgaacatgcccgaacatgcccgaaatgccccaagaagaagc 353
```

```
Db    294 GTTCTACGAATTTTCAAACGCCTCGTCCCGAAATGCCCGAAATGCCCCAAGAAGAAGC
Qy    354 agatgacgcggcggattgaacttcggttcggggcttcatcatcagcaaa 399
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    354 AGATGACGCGGCGGATTGAACTTCGGTTCGGGGCTTCATCATCAGCAAA 399

RESULT  3
NMA322491/c
LOCUS       NMA322491    311321 bp    DNA             BCT       30-MAR-2000
DEFINITION  Neisseria meningitidis serogroup A strain Z2491 complete genome;
            segment 3/7.
ACCESSION   AL162754 AL157959
VERSION     AL162754.2   GI:7379424
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 311321)
AUTHORS     Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
            Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
            Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
            Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
            Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
TITLE       Complete DNA sequence of a serogroup A strain of Neisseria
            meningitidis Z2491
JOURNAL     Nature 404 (6777), 502-506 (2000)
MEDLINE     20222556
REFERENCE   2  (bases 1 to 311321)
AUTHORS     Parkhill,J.
TITLE       Direct Submission
JOURNAL     Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
            sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of N. meningitidis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES             Location/Qualifiers
     source          1..311321
                     /organism="Neisseria meningitidis"
                     /strain="Z2491"
                     /db_xref="taxon:487"
                     /note="serogroup: A"
     RBS             115..120
     gene            129..587
                     /gene="NMA0698"
     CDS             129..587
                     /gene="NMA0698"
                     /note="NMA0698, possible periplasmic protein, len: 152 aa;
                     unknown, contains a probable N-terminal signal sequence
                     and lies within a region of unusually low GC content"
                     /codon_start=1
                     /transl_table=11
                     /product="putative periplasmic protein"
                     /protein_id="CAB83985.1"
                     /db_xref="GI:7379425"
                     /translation="MLTKLKILFLFLFVVFVLAINLLFFFSSDIESFGNYQFYYYID
                     KGWPANYILVMKDGNEGNFDKIISGLVLEYYKEDDNIYFSYIDGQFASDSCYYKPEI
                     LYGKIILNKHIINSMEKNNFLSEDKIMKGTRNWLADPKNKNIQTLD"
     gene            742..1077
                     /gene="NMA0699"
     CDS             742..1077
                     /gene="NMA0699"
                     /note="NMA0699, pseudogene, probable ABC transporter
                     protein, len: 336 bp; similar to C-termini of many ABC
                     transporters e.g. TR:Q46973 (EMBL:U47048), mtfB,
                     Escherichia coli microcin transport protein (707 aa),
                     fasta scores; E(): 2.6e-19, 53.2% identity in 111 aa
                     overlap. Contains PS00211 ABC transporters family
```

```
Db    294 GTTCTACGAATTTTCAAACGCCTCGTCCCGAAATGCCCGAAATGCCCCAAGAAGAAGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    354 agatgacgcggcggattgaacttcggttcggggcttcatcatcagcaaa 399
          |||||||||||||||||||||||||||||||||||||||||||||
Db    354 AGATGACGCGGCGGATTGAACTTCGGTTCGGGGCTTCATCATCAGCAAA 399
```

```
misc_feature    signature. NMA0686 may be the remainder of this
                pseudogene"
                /codon_start=1
                /pseudo
                /transl_table=11
                /product="putative ABC transporter protein (pseudogene)"
gene            868..912
                /gene="NMA0699"
                /note="PS00211 ABC transporters family signature"
                complement(1101..2327)
                /gene="NMA0700"
CDS             complement(1101..2327)
                /gene="NMA0700"
                /note="NMA0700, possible ribonuclease BN, len: 408 aa;
                similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
                coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
                2.3e-28, 34.3% identity in 268 aa overlap. Longer than rbn
                at the C-terminus. Also similar to TR:O85449
                (EMBL:AF067083) vitreoscilla sp. hypothetical protein (376
                aa), fasta scores: E(): 0, 45.0% identity in 333 aa
                overlap. Contains a region similar to NMA0677, fasta
                scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
                Contains hydrophobic, possible membrane-spanning regions,
                as does rbn"
                /codon_start=1
                /transl_table=11
                /product="putative ribonuclease BN"
                /protein_id="CAB83987.1"
                /db_xref="GI:7379426"
                /translation="MTFLQRLQGLADNKICAFAWFVVRRFDEERVPQAAASMTFTLL
                AIVPVLIVMVAVASIFPVFDRWSDSFVSFVNQTIVPQGADMVFDYINAFREQANRLTA
                IGSVMLVVTSLMLIRTIDNTFNRIWRVNSQRPWMMQFLVYWALLTFGPLSLGVGISFM
                VGSVQDAALASGAPQWSGALRTAATLTFMTLLLWGLYREVPNREFVPARQAFVGALATA
                FCLETARSLFTWYMGNFDGYRSIYGAFAAVPFFLIWLNLLWTLVLGGAVLTSSLSYWQ
                GEAFRRGFPDSRGRFDDVLKILLLLDAAQKEGKALPVQEFRRHINMGYDELGELLEKLA
                RHGYIYSGRGQWVLKTGADSIELNELFKLFVYRPLPVERDHVNQAVDAVMPCLQTLN
                MTLAEFDAQAKKQQQS"
misc_feature    1376..1385
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    complement(2417..2426)
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    2444..2453
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
gene            2486..2743
                /gene="NMA0701"
CDS             2486..2743
                /gene="NMA0701"
                /note="NMA0701, possible pseudogene, len: 258 bp; shows
                weak similarity to part of SW:WRBA_ECOLI (EMBL:M99166),
                wrbA, Escherichia coli Trp repressor binding protein (197
                aa), fasta scores: E(): 0.064, 29.3% identity in 82 aa
                overlap. Also similar to part of TR:O85450 (EMBL:AF067083)
                vitreoscilla sp. Trp repressor binding protein (fragment)
                (124 aa), fasta scores: E(): 6.6e-14, 55.6% identity in 83
                aa overlap"
                /codon_start=1
                /pseudo
                /transl_table=11
                /product="hypothetical protein NMA0701 (pseudogene)"
misc_feature    complement(2755..2764)
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    2778..2787
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
RBS             2861..2864
                /gene="NMA0702"
gene            2871..3530
                /gene="NMA0702"
CDS             2871..3530
                /gene="NMA0702"
                /note="NMA0702, len: 219 aa; similar to TR:Q52925
                (EMBL:Z50189), exsB, Rhizobium meliloti putative regulator
                of succinoglycan biosynthesis (not a transcriptional
                regulator) (234 aa), fasta scores: E(): 3.1e-15, 32.7%
                identity in 205 aa overlap. Also similar to many
                hbacterial ypothetical proteins e.g. SW:YBAX_HAEIN
                (EMBL:U32798), HI1191, Haemophilus influenzae hypothetical
                protein (196 aa), fasta scores: E(): 0, 79.2% identity in
                173 aa overlap"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein NMA0702"
                /protein_id="CAB83989.1"
                /db_xref="GI:7379427"
                /translation="MSNQKALVIFSGGQDSTTCLIQAIQTYGRENVQAITFQYGQRHA
                VELERARWIAQDLGVKQTVLDLSLMRQITHNALMDDTAAIETAENGVPNTFVDGRNAL
                FLLYAAIYAKGQGIRHIIAGVCETDFSGYPDCRDVFVKSMNVTLNLAMDYDFQIHTPL
                MYLTKAQTWALADEMGALDYIREQTHTCYNGIVGGCRECPSCILRERGLAEYLESKKA
                V"
misc_feature    3522..3531
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
RBS             3547..3550
gene            3558..4076
                /gene="NMA0703"
CDS             3558..4076
                /gene="NMA0703"
                /note="NMA0703, len: 172 aa: unknown"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein NMA0703"
                /protein_id="CAB83990.1"
                /db_xref="GI:7379428"
                /translation="MPKLHMFYLGGNAGRSNIEVHDIQFAVCDDYREAVPALKAAWFG
                DADRIHIDGWQVVEWADGYDIAVSETPKTKMPSENAPRLYFANVGGYRAGQLEAHAF
                GLFRAATPAEAKQKALQTLLTDSYVQQHKDNLKDVDNLLALDRIGNFHIRLTPNPHGK
                PAEIGFQGYLPI"
misc_feature    3782..3791
                /gene="NMA0703"
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    4083..4502
                /gene="NMA0704"
                /note="Pfam match to entry PF01242 PTPS, 6-pyruvoyl
                tetrahydropterin synthase, score 10.80, E-value 1.7e-06"
gene            4083..4505
                /gene="NMA0704"
CDS             4083..4505
                /gene="NMA0704"
                /note="NMA0704, len: 140 aa; similar to many hypothetical
                proteins e.g. SW:YB9Q_HAEIN (EMBL:U32798), HI1190,
                Haemophilus influenzae hypothetical protein (141 aa),
                fasta scores: E(): 5.1e-32, 58.0% identity in 138 aa
                overlap. Shows very weak similarity to eukaryotic
                6-pyruvoyl-tetrahydropterin synthases e.g. SW:PTPS_RAT
                (EMBL:M77850), pts, Rattus norvegicus 6-pyruvoyl
                tetrahydrobiopterin synthase precursor (EC 4.6.1.10) (144
                aa), fasta scores: E(): 0.0062, 27.3% identity in 154 aa
                overlap. Contains Pfam match to entry PF01242 PTPS,
                6-pyruvoyl tetrahydropterin synthase"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein NMA0704"
                /protein_id="CAB83991.1"
```

```
Query Match            12.5%;  Score 175;  DB 2;  Length 311321;
Best Local Similarity  99.6%;  Pred. No. 2.6e-84;
Matches 225;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy    174 cdaactggtcaaagcgaagcccggcagtcgtcaatattcaggcagcccccccgcg 233
          |||||||||| ||||||||| |||||||||||||||||||||||||||||||| ||
Db  10648 CCAACTGGTTCAAAGCGAAGGCCCGGCAGTCGTCAATATTCAGGCAGCCCCCCGCCG 10589
```

```
Qy    234  cacccaaaacggcagcgggcaatgccgaaaccgattccgaccgcttgccgacagcggaccc  293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10588  CACCCAAAACGGCAGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGGACCC  10529

Qy    294  gtcctacgaatttttcaaacgcctcgtcccgaacatgccgaaatcccccaagaagaagc  353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10528  GTTCTACGAATTTTTCAAACGCCTCGTCCCGAACATGCCCGAAATCCCCCAAGAAGAAGC  10469

Qy    354  agatgacgacggattgaacttcggttcggcttcatcatcagcaaa  399
           |||||||||||||||||||||||||||||||||||||||||||||
Db  10468  AGATGACGACGGCGGATTGAACTTCGGTTCGGCGGCTTCATCATCAGCAAA  10423
```

```
RESULT  4
AX011576
LOCUS       AX011576      1500 bp      DNA            BCT         18-SEP-2000
DEFINITION  Sequence 5 from Patent WO9955872.
ACCESSION   AX011576
VERSION     AX011576.1  GI:9998107
KEYWORDS    .
SOURCE      Neisseria meningitidis.
ORGANISM    Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1500)
AUTHORS     Ruelle,J.L.
TITLE       Basb013 dna and proteins from neisseria meningitidis
JOURNAL     Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES            Location/Qualifiers
     source          1..1500
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT    385 a    471 c    382 g    262 t
ORIGIN
```

```
Query Match             11.2%;  Score 156;  DB 2;  Length 1500;
Best Local Similarity   99.0%;  Pred. No. 7.3e-74;
Matches  306;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0:

Qy    991  cccgcagcccgagacgtgccggcctgcaggcggcgacatctgtcctcagcctcgac  1050
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    991  CCCGCAGCCCGCAGAAACGTGCCCGCCTGCAGGCGGCGGACATCGTCCTCAGCCTCCGAC  1050

Qy   1051  gggcgagaaatacgttcttccgcgacctccgtcatggtcggcgcgccattacgccggga  1110
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1051  GGCGGAGAAATACGTTCTTCCGCGGACCTTCCCGTTATGGTCGGCGCGCCATTACGCCGGGA  1110

Qy   1111  aagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacacaagccaagctg  1170
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1111  AAGAAGTCAGCCTCGGCGTATGGCGCAAAGGCGAAGAAATCACAATCAAAGTCAAGCTG  1170

Qy   1171  ggcaacgcgccgagcataccgggcgcatcatccaaaacagatgaagcccctacaccgaa  1230
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1171  GGCAACGCGCCGAGCATACCGGGCGCATCATCCAAAACAGATGAAGCCCCTACACCGAA  1230

Qy   1231  cagcaatccggtacgttctcggtcgaatccgcaggcattaccccttcagacacataccgac  1290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1231  CAGCAATCCGGTACGTTCTCGGTCGAATCCGCAGGCATTACCCCTTCAGACACCATACCGAC  1290

Qy   1291  agcagcggc  1299
           |||||||||
Db   1291  AGCAGCGGC  1299
```

```
RESULT  5
AE002409/c
LOCUS       AE002409    10531 bp    DNA            BCT         25-MAY-2000
DEFINITION  Neisseria meningitidis serogroup B strain MC58 section 51 of 206 of
            the complete genome.
ACCESSION   AE002409 AE002098
VERSION     AE002409.1  GI:7225757
```

```
KEYWORDS    .
SOURCE      Neisseria meningitidis MC58.
ORGANISM    Neisseria meningitidis MC58
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 10531)
AUTHORS     Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
TITLE       Complete genome sequence of Neisseria meningitidis serogroup B
            strain MC58
JOURNAL     Science 287 (5459), 1809-1815 (2000)
MEDLINE     20175755
PUBMED      10710307
REFERENCE   2  (bases 1 to 10531)
AUTHORS     Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
TITLE       Direct Submission
JOURNAL     Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
```

```
FEATURES            Location/Qualifiers
     source          1..10531
                     /organism="Neisseria meningitidis MC58"
                     /strain="MC58"
                     /db_xref="taxon:122586"
                     /note="serogroup: B"
     gene            complement(132..1631)
                     /gene="NMB0532"
                     /note="This region contains an authentic frame shift and
                     is not the result of a sequencing artifact; similar to
                     PID:1184684 PID:1220194 percent identity: 66.59;
                     identified by sequence similarity; putative:protease DO,
                     authentic frameshift"
     gene            complement(1769..2398)
                     /gene="NMB0533"
     CDS             complement(1769..2398)
                     /gene="NMB0533"
                     /note="similar to GB:J02857 SP:P20625 PID:146972 GB:U00096
                     PID:1742691 percent identity: 86.12; identified by
                     sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="endonuclease III"
                     /protein_id="AAF40962.1"
                     /db_xref="GI:7225758"
                     /translation="MNRHIRQEIFERFRAANPHPTELNFNSPFELLIAVLLSAQATD
                     VGVNKATAKLFPVADTPQAMLDLGLDGVMEYTKTIGLYKTKSKHIMQTCRILLEKYNG
                     EVPEDREALESLPGVGRKTANVVLNTAFGHPVMAVDTHIFRVSNRTKIAPGKDVREVE
                     DKLMRFIPKEFLMDAHHWLLHGRYTCKALKPQCCQTCIINDLCEYPAKA"
     gene            complement(2444..2857)
                     /gene="NMB0534"
     CDS             complement(2444..2857)
                     /gene="NMB0534"
                     /note="conserved hypothetical protein; identified by
                     Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF40963.1"
                     /db_xref="GI:7225759"
```

```
gene            /translation="MIPSYTYAALAFSAFTSATLLPGTSEARAFALFVHRFPEHAYGAL
                LCAGLANGLGSMVSYWMGRLLPSRKMPSEKTLNLMRRFGIWLLAFTWLPVVGDALPLA
                AGWLRLNPWTSGLMLVIGKTARYAFILWGMQYYAA"
                3390. .4670
                /gene="NMB0535"
CDS             3390. .4670
                /gene="NMB0535"
                /note="similar to PID:1171339 percent identity: 71.89;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="glucose/galactose transporter"
                /protein_id="AAF40964.1"
                /db_xref="GI:7225760"
                /translation="MSISIFLKRVIFRHHERQIMSAQSQNNHTSPLVVLTLFFMMGF
                TTCMNDLIPHLKEIFDLSYVQAMLIQPCFFTAYAVMSIPMGAFVGKVGYKNGVIGGF
                LETAVGCLLFYPAAGSHSYAVFLGALFILASGVTLLQVAGNPVVTLLAKPGKFSATLT
                LVQAFNALGTTIAPQIGAFLILADATQTVSKAEQISSVQIPYLGLAGLLILAVFVKM
                IRLPDARKIAAEESAHNHDGKTSVWQYKHLVFGTAGIFCYVAGAVSIGSLMVNVLGYL
                KGLDHASAAHYLSFYWGGAMVGRFLGSAVMAKFAPNRYLAFNASAAVVLLAVAMATGS
                GNADVAMWSLLAIGFFNSIMFPTIFSLATVKGLGKFTNAASGVLCTAIVGGAVVPVVQG
                WVADTYTLMSSFVVSVICYLYIVPFAVVCYRADK"
gene            4961. .6340
                /gene="NMB0536"
CDS             4961. .6340
                /gene="NMB0536"
                /note="similar to GB:M73530 SP:P27611 PID:143245
                PID:2209268 percent identity: 65.57; identified by
                sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="Na+/H+ antiporter"
                /protein_id="AAF40965.1"
                /db_xref="GI:7225761"
                /translation="MFAFKSLLDMPRGEALAVVVALIAAMGYTIISLEWLPHMSIIAA
                IVVLILYGLARGLKYNDMQQGMIGALNQGMGAIYLFFIGLMVSALMMSGAIPTLMYY
                GFGLISPTYPYFSSFALCSVVSIGSSLTCATVGVAFMGMAAAFQADMAMTAGAIV
                SGAFFGDKMSPLSDTTGISASIVGIDLFEHIKNMYTTIPAWLISAALMLLLPNVAA
                QDLNSVESFRSQLEATGLVHGYSLIPFALLVILALRINAVVAMLFTVMAVAVTYLH
                STPDLRQLGAWFYGGYKLEGEAFKDVVKLLISRGGLESMFFTQTIVILGMSLGGLLFAL
                GVIPSLLEAIRTFLTNAGRATFSVAMTSVGVNFLIGEQYLSILLSGETFKPVYDKLGL
                HSRNLSRTLEDAGTVINPLVPWSVCGVFISHALGVPWEYLPYAFFCYLSLALTLLFG
                WTGLTLSKK"
gene            complement(6469. .7293)
                /gene="NMB0537"
CDS             complement(6469. .7293)
                /gene="NMB0537"
                /note="conserved hypothetical protein; identified by
                Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAF40966.1"
                /db_xref="GI:7225762"
                /translation="MDGAQPKTNFFERLIARLAREPDSAEDVLNLLRQAHEQEVFDAD
                TLLRLERKVLDFSDLEVRDAMITRSRMNVLKENDSIERITAYVIDTAHSRFPVIGEDKD
                EVLGILHAKDLLKYMFNPEQFHLKSILRPAVFVPEGKSLTALLKEFRQRNHMAIVID
                EYGGTSGLVTPEDIIEQIVGEIEDEFDEDDSADNIHAVSSERWRIHAATEIEDINTFF
                GTEYSSEADTIGGLVIQELGHIPVRGEKVLIGGLQFTVARADNRRLHTLMATRVK"
gene            complement(7301. .7810)
                /gene="NMB0538"
CDS             complement(7301. .7810)
                /gene="NMB0538"
                /note="conserved hypothetical protein; identified by
                Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAF40967.1"
                /db_xref="GI:7225763"
                /translation="MKRFKKYPFLTLQRQRFHLNFENASSAAGIPAERDFYRWAWSAL
                KNEYRRADISLILLDEEERAYNRDYRGKDYATNVLSFALNEGEILPCGVSEKLYGDL
                IICPQVVLKEAAEQGKTPEQHFAHLTIHGTLHLMGYDHIEDDEAEIMEAEEIRLMRAA
```

```
                GFPNPYQED"
gene            7929. .8864
                /gene="NMB0539"
CDS             7929. .8864
                /gene="NMB0539"
                /note="similar to PID:151255 PID:496214 SP:Q60169 percent
                identity: 72.96; identified by sequence similarity;
                putative"
                /codon_start=1
                /transl_table=11
                /product="porphobilinogen deaminase"
                /protein_id="AAF40968.1"
                /db_xref="GI:7225764"
                /translation="MNPKKLVIASRESLLAMWQAKHIQGRLKALYPDCEVEILGMTTR
                GDQILDKTLSKVGGKGLFVKELEQALYDGRADLAVHSIKDVPMDLPEGFALAAIGERA
                NPFDAFVSNQYTRLEEMPEGAVVGTSSLRREQLRARYPHLLIKPLRGNVQTRLSKLD
                NGEYDAIILAAGLQRIKLDGRIRMLSESDSLPAAGQGALGIEIAAHREDLYEVLKP
                LNHGVTNACVTAERALARALGGSCQVPLAAYCTEENGLLTLRGLVGHPDGSVVLRADA
                QAPAEYADALGRAVAKKLADDGARELIGAVLNTEN"
gene            complement(9268. .10461)
                /gene="NMB0540"
CDS             complement(9268. .10461)
                /gene="NMB0540"
                /note="similar to GB:L42023 SP:P44425 PID:1007884
                PID:1221762 PID:1205848 percent identity: 84.60;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="aspartate aminotransferase"
                /protein_id="AAF40969.1"
                /db_xref="GI:7225765"
                /translation="MFFKHIEAPADPILGLEAFKAETRPEKVNLGIGVYKDASGAT
                PLVKAVKEAEKRLLESETTKNYLTIDGVADYNAQTQILLFGKDHEIIASRAKTAQSL
                GGTGALRIAAEFAKRQLNAQTIWISNP'WPNHNAIAKAVGIQDKPYRYDAAKHGLDW
                DGMIEDLSQAQKGDIVLLHGCCHNPTGIDPTPEQWETLAKLSAEKGWLPLFDAYQGF
                GNLEEDAYGLRVFLKHNTELLIASSYSKNFGMYNERVGAFTLVAEDEETAARAHSQV
                KTIIRTLYSNPASHGANTIALVLKNDDLKAQWIAELDEMRGRIKAMRQKFVGLLKAKG
                ASQNFDFIIKQNGMFSFSGLTPEQVDRLKNEFAIYAVRSGRINVAGITDNNIDYLCES
                IVKVL"
```

```
BASE COUNT      2151 a    2498 t   2982 g   2900 t
ORIGIN
```

```
Query Match        11.2%;   Score 156;   Length 10531;
Best Local Similarity 99.08%; Pred. No. 7.1e-74;
Matches 306;  Conservative  0;  Mismatches  3;  Indels  3;  Gaps  0;

Qy    991  cccggcagcccgcagaacgtgccggctgcaggcgggcgacatcgtcctcagcctcgac 1050
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    641  CCCGGCAGCCCGCAGAACGTGCCGGCCTGCAGGCGGGCGACATCGTCCTCAGCCTCGAC 582

Qy   1051  ggcggagaaatacgttcttccggcgacctccgtcatgtcggcggccattacggcgggga 1110
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    581  GGCGGAGAAATACGTTCTTCCGGCGACCTCCGTTATGGTCGGCGGCCATTACGCCGGGA 522

Qy   1111  aaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacacatcaaagccaagctg 1170
           |||||||||||||||| ||||||||||||||||||||| ||||||| |||||
Db    521  AAAGAAGTCAGCCTCGGCGTATGGCGCAAAGGCGAAGAAATCACAATCAAAGTCAAGCTG 462

Qy   1171  ggcaacgcccgcgacataccggcgcatcctcaaaacagatgaagcccccctacaccggaa 1230
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461  GGCAACGCCCGCGACATACCGGCGCATCATCAAAACAGATGAAGCCCCTACACCGAA 402

Qy   1231  cagcaatccggtacgttctcggtcgaatccgcaggcattaccctcagacacataccgac 1290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401  CAGCAATCCGGTACGTTCTCCGGTCGAATCCGCAGGCATTACCCTTCAGACACATACCGAC 342

Qy   1291  agcagcggc 1299
           |||||||||
Db    341  AGCAGCGGC 333
```

```
RESULT          6
```

Fri Mar 23 07:26:15 2001   us-09-388-090-3.oligo25.rge

Search completed: March 22, 2001, 02:51:11
Job time: 4531 sec

```
AX011578
LOCUS       AX011578     1110 bp     DNA              BCT       18-SEP-2000
DEFINITION  Sequence 7 from Patent WO9955872.
ACCESSION   AX011578
VERSION     AX011578.1  GI:9998108
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   Ruelle,J.L.
  TITLE     Basb013 dna and proteins from neisseria meningitidis
  JOURNAL   Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1..1110
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      280 a    350 c    279 g    201 t
ORIGIN

Query Match             8.1%;   Score 113;  DB 2;  Length 1110;
Best Local Similarity  100.0%;   Pred. No. 2.8e-50;
Matches 113;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    439 atggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgcccaactcatcggt 498
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    439 ATGGCAGTATCAAAGTCCTGCTCAACGACAAGCGGCAATATACCGCCCAAACTCATCGGT 498

Qy    499 tcggatgtccaatccgatgtcgcccttctgaaaatcgacgcgcaacggaagagct 551
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    499 TCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCGCAACGGAAGAGCT 551


RESULT 7
AX011580
LOCUS       AX011580     390 bp     DNA               BCT       18-SEP-2000
DEFINITION  Sequence 9 from Patent WO9955872.
ACCESSION   AX011580
VERSION     AX011580.1  GI:9998109
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Ruelle,J.L.
  TITLE     Basb013 dna and proteins from neisseria meningitidis
  JOURNAL   Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1..390
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      106 a    120 c    102 g     62 t
ORIGIN

Query Match             7.8%;   Score 109;  DB 2;  Length 390;
Best Local Similarity  100.0%;   Pred. No. 4.5e-48;
Matches 109;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1191 cggcgcatcatccaaaacagatgagcccctacaccgaacagcaatccggtacgttctc 1250
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 CGGCGCCATCATCCAAAACAGATGAGCCCCTACACCGAACAGCAATCCGGTACGTTCTC 140

Qy   1251 ggtcgaatccgcaggcattaccccttcagacacatacgacagcagcggc 1299
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GGTCGAATCCGCAGGCATTACCCCTTCAGACACATACCGACAGCAGCGGC 189
```

us-09-388-090-3.oligo25.rge

Fri Mar 23 07:26:15 2001

```
                    GenCore version 4.5
            Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        March 22, 2001, 01:37:50 ; Search time 90.77 Seconds
                                  (without alignments)
                                  5773.377 Million cell updates/sec

Title:         US-09-388-090-3
Perfect score: 1395
Sequence:      1 gtgttcaaaaataccaata........tcagggcaagtcccgtcaa 1395

Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0

Searched:      480022 seqs, 187831343 residues

Word size :    25

Total number of hits satisfying chosen parameters:     18

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :     N_Geneseq_36:*
                1: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
                2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
                3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
                4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
                5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
                6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
                7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
                8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
                9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
               10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
               11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
               12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
               13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
               14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
               15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
               16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
               17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
               18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
               19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
               20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
               21: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	1395	100.0	1395	21	Z93414	NGSP polypeptide c
2	1344	96.3	1395	21	Z51538	Neisseria meningit
3	1323	94.8	1326	21	Z51539	Neisseria meningit
4	1248	89.5	1500	21	Z54510	Neisseria gonorrhe
5	1242	89.0	1242	21	Z93415	NGSP polypeptide c
6	175	12.5	1347	21	Z51533	Neisseria meningit
7	175	12.5	1500	21	Z33305	Neisseria meningit
8	175	12.5	1500	21	Z33306	Neisseria meningit
9	156	11.2	1500	21	Z54511	Neisseria meningit
10	156	11.2	1500	21	Z33307	Neisseria meningit
11	153	11.0	153	21	Z93416	NGSP coding sequen
12	138	9.9	1500	21	Z54512	Neisseria meningit
13	113	8.1	1110	21	Z33308	Neisseria meningit
14	109	7.8	390	21	Z33309	Neisseria meningit
15	38	2.7	153	21	Z51542	Neisseria meningit
c 16	33	2.4	69	21	Z93413	Primer used for am
c 17	30	2.2	30	21	Z51537	3' PCR primer to am
18	30	2.2	46	21	Z93412	Primer used for am

ALIGNMENTS

```
RESULT  1
Z93414
ID   Z93414 standard; DNA; 1395 BP.
XX
AC   Z93414;
XX
DT   24-JUL-2000 (first entry)
XX
DE   NGSP polypeptide coding sequence.
XX
KW   NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW   cellular matrix; adherence domain; ligand; detection; diagnosis;
KW   screening; probe; primer; prophylaxis; therapy; ss.
XX
OS   Neisseria gonorrhoeae.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1395
FT                  /*tag= a
FT                  /product= NGSP polypeptide
XX
PN   WO200012133-A1.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US20070.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ, Harris AM;
XX
DR   WPI; 2000-237782/20.
DR   P-PSDB; Y83150.
XX
PT   Non-cytosolic NGSP polypeptide and polynucleotide sequence from
PT   Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT   infections
XX
PS   Claim 38; Page 60-61; 68pp; English.
XX
CC   The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC   Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC   domains for extracellular matrix proteins. Using the NGSP polypeptide
CC   as a vaccine produces antibodies which inhibit binding of N.
CC   gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC   subsequent invasion. The NGSP polypeptide and its peptide fragments
CC   can be used to immunise an animal and produce an immune response.
CC   They can also be used as ligands to detect antibodies elicited in
CC   response to Neisseria infections and also as antigens or immunogens
CC   for inducing Neisseria-specific antibodies which are useful in
CC   immunoassays to detect Neisseria in biological specimens. Nucleotides
CC   encoding NGSP or its fragments can be used as probes to identify
CC   Neisseria in biological specimens by hybridization or polymerase
CC   chain reaction amplification. The NGSP polypeptide can also be used
CC   in screening assays to identify agents and compounds which useful as
CC   diagnostic, prophylactic or therapeutic agents against Neisseria
CC   infection.
XX
SQ   Sequence 1395 BP; 359 A; 438 C; 353 G; 245 T; 0 other;
```

Query Match   100.0%; Score 1395; DB 21; Length 1395;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1395; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 gtgttcaaaaataccaatacttcgcttggcgcactgtgccgccttggccaggc   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 gtgttcaaaaataccaatacttcgcttggcgcactgtgccgccttggccaggc   60

Qy   61 tgcgaaaaggcaggcagcttttcgtgcggacaaaaagaagcatcctcgtagaacgc  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 tgcgaaaaggcaggcagcttttcgtgcggacaaaaagaagcatcctcgtagaacgc  120

Qy  121 atcgaacaccaccaagacgacgacgcagtgtcagtatgctgtcgccgactttgccaactg  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 atcgaacaccaccaagacgacgacgcagtgtcagtatgctgtcgccgactttgccaactg  180

Qy  181 gttcaaagcgaaggcccggaccgtcaatattcaggcagcgccccccgccgcaccaa  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 gttcaaagcgaaggcccggaccgtcaatattcaggcagcgccccccgccgcaccaa  240

Qy  241 aacggcagcggcaatgccgaaaccgattcgaccgcttgccgacagcgaccgttctac  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 aacggcagcggcaatgccgaaaccgattcgaccgcttgccgacagcgaccgttctac  300

Qy  301 gaattttcaaacgcctcgtccgaacatgcccgaaatccccaagaagaagcagatgac  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 gaattttcaaacgcctcgtccgaacatgcccgaaatccccaagaagaagcagatgac  360

Qy  361 ggccgattgaacttcggttcgggttcggttcggtccgaacacacggctacatcctgaccaat  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ggccgattgaacttcggttcgggttcggttcggttcggtccgaacacacggctacatcctgaccaat  420

Qy  421 acccacgtcgttgccggtatgggcagtatcaaatcctctcaacgacaagcgcaaatat  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 acccacgtcgttgccggtatgggcagtatcaaatcctctcaacgacaagcgcaaatat  480

Qy  481 acggccaaactcaatcggttcggatgtcgtccaatccgatgtcgcccttctgaaatcgacgca  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 acggccaaactcaatcggttcggatgtcgtccaatccgatgtcgcccttctgaaatcgacgca  540

Qy  541 acggaagagctaccgtcgctgcaaatcggcaatcccaaaatttgaaacgggcgaatgg  600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 acggaagagctaccgtcgctgcaaatcggcaatcccaaaatttgaaacgggcgaatgg  600

Qy  601 gtcgctgcatcggcgcgcgccccttcggctttgacaacagcgtgacccgccgcatccgtgtcc  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 gtcgctgcatcggcgcgcgccccttcggctttgacaacagcgtgacccgccgcatccgtgtcc  660

Qy  661 gccaaaggcagaagccgaagcctgccaacgcagaaagctacaccccttcatccaaaccgacgttgcc  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 gccaaaggcagaagccgaagcctgccaacgcagaaagctacaccccttcatccaaaccgacgttgcc  720

Qy  721 atcaatccgggcaattccggcccgctgttcaacttaaaggacaggtcgtcgcatc  780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 atcaatccgggcaattccggcccgctgttcaacttaaaggacaggtcgtcgcatc  780

Qy  781 aattcgcaaatatacagcgcggattcatgggcatctccttgccatcccgatt  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 aattcgcaaatatacagcgcggattcatgggcatctccttgccatcccgatt  840

Qy  841 gacgttgccatgaatgtgccgacagctgaaaacaccggcaagtccaacgcggacaa  900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 gacgttgccatgaatgtgccgacagctgaaaacaccggcaagtccaacgcggacaa  900

Qy  901 ctgggctgattattcaggaagtatatcctacggtttggcacagtcgttcggtctggataaa  960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 ctgggctgattattcaggaagtatatcctacggtttggcacagtcgttcggtctggataaa  960

Qy  961 gccagcggcgcattgattgccaaatcctccccgcagacagccccgcagaacgtgcccggcctg 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 gccagcggcgcattgattgccaaatcctccccgcagacagccccgcagaacgtgcccggcctg 1020
```

```
Qy 1021 cagcggggcgacatcgtcctccagcctcgacgcgcggagaaatacgtcttccggcgcacctt 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 cagcggggcgacatcgtcctccagcctcgacgcgcggagaaatacgtcttccggcgcacctt 1080

Qy 1081 cccgtcatgtcggcgccattacgccgggaaaagagtcagcctcggcgtatggccgcaa 1140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 cccgtcatgtcggcgccattacgccgggaaaagagtcagcctcggcgtatggccgcaa 1140

Qy 1141 ggcgaagaaatcacacatcaaagccaagctggcaacgccgcagcacatccggccgcatca 1200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 ggcgaagaaatcacacatcaaagccaagctggcaacgccgcagcacatccggccgcatca 1200

Qy 1201 tccaaaacagatgaagcccctacacagcaacagcaatccggtacgttctcggtcgaatcc 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 tccaaaacagatgaagcccctacacagcaacagcaatccggtacgttctcggtcgaatcc 1260

Qy 1261 gcaggcattaccctcagacacataccgacagcagcggcaaacacctcgtcgtcacgg 1320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 gcaggcattaccctcagacacataccgacagcagcggcaaacacctcgtcgtacgg 1320

Qy 1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380

Qy 1381 gcaagtcccgtcaa 1395
        ||||||||||||||
Db 1381 gcaagtcccgtcaa 1395
```

```
RESULT   2
Z51538
ID   Z51538 standard; DNA; 1395 BP.
XX
AC   Z51538;
XX
DT   03-JUL-2000 (first entry)
XX
DE   Neisseria meningitidis NMASP protein-2 encoding DNA.
XX
KW   NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW   anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW   Neisserial infection; meningitidis; septicaemia; ds.
XX
OS   Neisseria meningitidis.
XX
FH   Key        Location/Qualifiers
FT   CDS        1..1395
FT              /tag= a
FT              /product= "NMASP protein"
FT              /note= "Does not include stop codon"
FT              /transl_except= (pos:22..24, aa:Leu)
FT              /transl_except= (pos:49..51, aa:Ser)
FT              /transl_except= (pos:64..66, aa:Asp)
FT              /transl_except= (pos:124..126, aa:Lys)
FT              /transl_except= (pos:172..174, aa:Val)
FT              /transl_except= (pos:250..252, aa:Ser)
FT              /transl_except= (pos:400..402, aa:Asp)
FT              /transl_except= (pos:433..435, aa:Thr)
FT              /transl_except= (pos:580..582, aa:Asp)
FT              /transl_except= (pos:649..654, aa:Gly)
FT              /transl_except= (pos:964..966, aa:Gly)
FT              /transl_except= (pos:1021..1023, aa:Arg)
FT              /transl_except= (pos:1162..1164, aa:Val)
FT              /transl_except= (pos:1189..1191, aa:Ile)
FT              /transl_except= (pos:1300..1302, aa:Gly)
FT              /transl_except= (pos:1354..1356, aa:Arg)
FT              /transl_except= (pos:1378..1395, aa: gqvpvndeagfrkamdkagk
FT              nvpllimrgntlfialnlq)
FT              /partial
XX
PN   WO200012535-A2.
XX
PD   09-MAR-2000.
```

```
XX
PF   01-SEP-1999;   99WO-US19663.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ, Harris AM;
XX
DR   WPI; 2000-256581/22.
DR   P-PSDB; Y70413.
XX
PT   Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT   antibodies, useful in vaccines against infection -
XX
PS   Claim 41; Page 69; 75pp; English.
XX
CC   The present sequence is a DNA encoding Neisseria meningitidis NMASP
CC   protein. NMASP is a non-cytosolic protein, with antibacterial and
CC   antiinflammatory activity. It shows sequence similarity to E. coli
CC   DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC   antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC   anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC   may be used for diagnosis, therapy or prophylaxis of Neisserial
CC   infections such as, bacterial meningitidis and septicaemia.
XX
SQ   Sequence 1395 BP; 358 A; 439 C; 353 G; 245 T; 0 other;

Query Match          96.3%;    Score 1344;    DB 21;    Length 1395;
Best Local Similarity 99.9%;    Pred. No. 0;
Matches 1394;    Conservative   0;    Mismatches   0;    Indels   1;    Gaps   0;

Qy    1  gtgttcaaaaataccaatacttcggttcggcggcactgtgtgccgccttgctggcaggc   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  gtgttcaaaaataccaatacttcggttcggcggcactgtgtgccgccttgctggcaggc   60

Qy   61  tgcgaaaggcagcagctgtttttcggtgcggacaaaagaagcatcctccgtagaacgc  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  tgcgaaaggcagcagctgtttttcggtgcggacaaaagaagcatcctccgtagaacgc  120

Qy  121  atcgaacacacacaaagacgacgcagtgtcagtatgctgctgctccgacttgcccaactg  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  atcgaacacacacaaagacgacgcagtgtcagtatgctgctgctccgacttgcccaactg  180

Qy  181  gttcaaagcggaggccggccgcagtcgtcaatattcaggcagcacatccgccccgcgccaccaa  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  gttcaaagcggaggccggccgcagtcgtcaatattcaggcagcacatccgccccgcgccaccaa  240

Qy  241  aacggcagcagccggcaatgccgaaaccgattccgacctcgccgacagcagccgttctac  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  aacggcagcagccggcaatgccgaaaccgattccgacctcgccgacagcagccgttctac  300

Qy  301  gaattttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  gaattttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac  360

Qy  361  ggcggattgaacttcgttcgggcttcatcatcagcaaaaacggctacatcctgaccaat  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  ggcggattgaacttcgttcgggcttcatcatcagcaaaaacggctacatcctgaccaat  420

Qy  421  acccacgtcgttgccgtatggcagtatcaagtcctgctcaacgacaagcgcgaatat  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  acccacgtcgttgccgtatggcagtatcaagtcctgctcaacgacaagcgcgaatat  480

Qy  481  acggccaaactcatcggttcggatgtgccaatcgatgcgcccttgaaatcgacgca  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  acggccaaactcatcggttcggatgtgccaatcgatgcgcccttgaaatcgacgca  540

Qy  541  acggaagagctaccgtcgtcaaatcggcaatcccaaaatttgaaccgggcgaatgg  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  acggaagagctaccgtcgtcaaatcggcaatcccaaaatttgaaccgggcgaatgg  600

Qy  601  gtcgctgccatcggcgcggccctccggcttgacaacagcgctgagacgcgcatcgtgtcc  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  gtcgctgccatcggcgcggccctccggcttgacaacagcgctgagacgcgcatcgtgtcc  660

Qy  661  gccaaggcagagctgccaacgaagctacacccctcatccaaacgacgttgcc  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  gccaaggcagagctgccaacgaagctacacccctcatccaaacgacgttgcc  720

Qy  721  atcaatccggcaattccggcggccccgcgttcaacttaaaggacaggtcgtcggcatc  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  atcaatccggcaattccggcggccccgcgttcaacttaaaggacaggtcgtcggcatc  780

Qy  781  aattcgcaaatatacagccgcagcggcggcggattcatggcatcccttgcatcccgatt  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  aattcgcaaatatacagccgcagcggcggcggattcatggcatcccttgcatcccgatt  840

Qy  841  gacgttgccatgaatgtcgccgacagcagcgtgaaaaacaccgcgcaagtccaacgcggacaa  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  gacgttgccatgaatgtcgccgacagcagcgtgaaaaacaccgcgcaagtccaacgcggacaa  900

Qy  901  ctggcggtgattattcaggaagaagtatcctacgttggcagtgcgtcgttcggtctggataaa  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  ctggcggtgattattcaggaagaagtatcctacgttggcagtgcgtcgttcggtctggataaa  960

Qy  961  gccagcggcgcattgattgccaaaatcttcccggacgccgcagacagtgccggcctg 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  gccagcggcgcattgattgccaaaatcttcccggacgccgcagacagtgccggcctg 1020

Qy 1021  caggcgggcgacacgtcctctcagcctcgacgcggcggagaaatacgttcttccggcgaccttt 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  caggcgggcgacacgtcctctcagcctcgacgcggcggagaaatacgttcttccggcgacctt 1080

Qy 1081  ccgtcatgtcggcgcgcattacggccggaaaaagagagtcagcctcggcgtatggcgcaa 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  ccgtcatgtcggcgcgcattacggccggaaaaagagagtcagcctcggcgtatggcgcaa 1140

Qy 1141  ggcgaagaaatcacaatcaaagccaagctggcacgccgccgagcatatcaccggcgcatca 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  ggcgaagaaatcacaatcaaagccaagctggcacgccgccgagcatatcaccggcgcatca 1200

Qy 1201  tccaaaacagatgaagcccctcaccgacaacaatccggtacgttctggtcggtcgaatcc 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  tccaaaacagatgaagcccctcaccgacaacaatccggtacgttctggtcggtcgaatcc 1260

Qy 1261  gcaggcattaccctcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  gcaggcattaccctcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320

Qy 1321  gtttccgacggcggagaacgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321  gtttccgacggcggagaacgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380

Qy 1381  gcaagtcccgtcaa 1395
         ||||||||||||||
Db 1381  gcaagtcccgtcaa 1395

RESULT   3
251539
ID   251539    standard; DNA; 1326 BP.
XX
AC   251539;
XX
DT   03-JUL-2000  (first entry)
XX
DE   Neisseria meningitidis NMASP protein-3 encoding DNA.
XX
KW   NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW   anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW   Neisserial infection; meningitidis; septicaemia; ds.
XX
OS   Neisseria meningitidis.
XX
```

```
Qy  253  aatgccgaaaccgattccgaccggttgccgacagcgaccgttctacgaattttcaaa  312
         |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  184  aatgccgaaaccgatccgaccggttgccgacagcgaccgttctacgaattttcaaa   243

Qy  313  cgcctcgtcccgaacatgcccgaaatcccccagaagaagcagatgacggcggattgaac  372
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244  cgcctcgtcccgaacatgcccgaaatcccccagaagaagcagatgacggcggattgaac  303

Qy  373  ttcggttcgggcttcatcatcagcaaaacgctacatcctgaccaataccccgtcgtt  432
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  304  ttcggttcgggcttcatcatcagcaaaacgctacatcctgaccaataccccgtcgtt  363

Qy  433  gccgtatgggcagtacaagtcctgctcaacgacaagcgaatataccgccaaactc  492
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364  gccgtatgggcagtacaagtcctgctcaacgacaagcgaatataccgccaaactc  423

Qy  493  atcggttcgatgtccaatcgatcgcgcctctgaaaatcgacgcacggaagagcta  552
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  424  atcggttcgatgtccaatcgatcgcgcctctgaaaatcgacgcacggaagagcta  483

Qy  553  ccgtcgtcaaaatcggcaatcccaaaaattgaaaccgggcgaatggtgtcgctgccatc  612
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484  ccgtcgtcaaaatcggcaatcccaaaaattgaaaccgggcgaatgggtcgctgccatc  543

Qy  613  ggcgcgcccttcggcttgcaacagcgtgaccgccggcacgtgtccgcaaaggcaga  672
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544  ggcgcgcccttcggcttgcaacagcgtgaccgccggcacgtgtccgcaaaggcaga  603

Qy  673  agcctgccaacgaagtcacaccttcatcaaccgacgttgcatcaatccggc  732
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  604  agcctgccaacgaagtcacaccttcatcaaccgacgttgccatcaatccggc  663

Qy  733  aattccggcggcccgtgtcaacttaaaagacagtcgtcggcatcaattcgcaaata  792
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  664  aattccggcggcccgctgtcaacttaaaagacagtcgtcggcatcaatccgcaaata  723

Qy  793  tacagccgcagcgcggcggcggcggatcatgggcatcttgccatccgattgacggttgccatg  852
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  724  tacagccgcagcgcggcggcggatcatgggcatcttgccatccgattgacggtgccatg  783

Qy  853  aatgtcgcgcaacagtcgaaaaacacggcaaagtcaacgcggacaactgggcgtgatt  912
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  784  aatgtcgcgcaacagtcgaaaaacacggcaaagtcaacgcggacaactgggcgtgatt  843

Qy  913  attcaggaagtatcccacggttggcacagtcgtcggtctggataaagccagcggcggcca  972
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  844  attcaggaagtatcccacggttggcacagtcgtcggtctggataaagccagcggcggcca  903

Qy  973  ttgattgccaaaatccttccggcagcccgcagaacgtgccggcctgcaggcggcgggcgac  1032
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  904  ttgattgccaaaatccttccggcagcccgcagaacgtgccggcctgcaggcgggcggcgac  963

Qy  1033 atcgtcctcagcctcgacggcggagaaatacgttcttccgcgtatggcgaaggcggaagaatc  1092
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  964  atcgtcctcagcctcgacggcggagaaatacgttcttccgcgtatggcgaaggcggaagaatc  1023

Qy  1093 ggcgccattacgccggaaaagaagtcagcctcggcgtatggcgcatccaaaacagat  1152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1024 ggcgccattacgccggaaaagaagtcagcctcggcgtatggcgcatccaaaacagat  1083

Qy  1153 acaatcaaagccaacggcatcggccggcagaacgcgcatccaaaacacacc  1212
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1084 acaatcaaagccaacggcatcggccgagcagaacgcgcatcatccaaaacacacc  1143

Qy  1213 gaagcccctacaccggcgaacagcaatccggtcgttccggtcgaatccggtcgatcattacc  1272
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1144 gaagcccccacaccggcgaacagcaatccggtcgttccggtcgaatccggtcgcattacc  1203

Qy  1273 cttcagacacatacccgacacaccggcaaacctcgtgtcgtacggttccgacgcg  1332
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1204 cttcagacacatacccgacacaccggcaaacacctcgtcgtacgggttccgacgcg  1263

Qy  1333 gcagaacggcgaggcttaaggcacggcgacgaaatcctagccgtcaggcgcaagtcccgt  1392
```

```
FH  Key             Location/Qualifiers
FT  CDS             1..1326
FT                  /*tag= a
FT                  /product= "NMASP protein"
FT                  /note= "Does not include stop codon"
FT                  /transl_except= (pos:55..57, aa:Lys)
FT                  /transl_except= (pos:103..105, aa:Val)
FT                  /transl_except= (pos:181..183, aa:Ser)
FT                  /transl_except= (pos:331..333, aa:Asp)
FT                  /transl_except= (pos:364..366, aa:Thr)
FT                  /transl_except= (pos:511..513, aa:Asp)
FT                  /transl_except= (pos:580..585, aa:Gly)
FT                  /transl_except= (pos:895..897, aa:Gly)
FT                  /transl_except= (pos:952..954, aa:Arg)
FT                  /transl_except= (pos:1093..1095, aa:Val)
FT                  /transl_except= (pos:1120..1122, aa:Ile)
FT                  /transl_except= (pos:1231..1233, aa:Gly)
FT                  /transl_except= (pos:1285..1287, aa:Arg)
FT                  /transl_except= (pos:1309..1311, aa:Gly)
FT                  /transl_except= (pos:1312..1314, aa:Gln)
FT                  /transl_except= (pos:1315..1317, aa:Val)
FT                  /transl_except= (pos:1321..1326, aa: vndeagfrkamdkagknvpl
FT                  limrrgntlfialnlq)
FT                  /partial
XX
PN  WO20001235-A2.
XX
PD  09-MAR-2000.
XX
PF  01-SEP-1999; 99WO-US19963.
XX
PR  01-SEP-1998; 98US-0098685.
XX
PA  (ANTE-) ANTEX BIOLOGICS INC.
XX
PI  Jackson WJ, Harris AM;
XX
DR  WPI: 2000-256581/22.
DR  P-PSDB: Y70414.
XX
PT  Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT  antibodies, useful in vaccines against infection -
PS  Claim 41; Page 73; 75pp; English.
XX
CC  The present sequence is a DNA encoding Neisseria meningitidis NMASP
CC  protein. NMASP is a non-cytosolic protein, with antibacterial and
CC  antiinflammatory activity. It shows sequence similarity to E. coli
CC  DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC  antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC  anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC  may be used for diagnosis, therapy or prophylaxis of Neisserial
CC  infections such as, bacterial meningitidis and septicaemia.
XX
SQ  Sequence 1326 BP; 342 A; 422 C; 334 G; 228 T; 0 other;

Query Match          94.8%;  Score 1323;  DB 21;  Length 1326;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1323;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   73  ggcagctttcggtgcggacaaaaggaagcatccttcgtgccgacttgccccactttcaggaacacacacc  132
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4  ggcagctttcggtgcggacaaaaggaagcatccttcgtgccgacttgccccactttcaggaacacacacc   63

Qy  133  aaagacgacgcagtcagtatgctgctgtccgacttgccccactttgccactggttcaaagcgaa  192
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64  aaagacgacgcagtcagtatgctgctgtccgacttgccccactttgccactggttcaaagcgaa  123

Qy  193  ggcccggcagtcgtcaatattcaggcagcagcccgcagcccccgcgcaccccaaaacggcagcggc  252
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124  ggcccggcagtcgtcaatattcaggcagcagcccgcagcccccgcgcaccccaaaacggcagcggc  183
```

Matches 1298;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   0:

```
Qy      56  caggctgcgaaagcaggcaggcagcttttcggtgcggacaaaaagaaggagcatcctcgtag  115
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56  caggctgcgaaagcaggcaggcagcttttcggtgcggacaaaaagaaggagcatcctcgtag  115

Qy     116  aacgcatcgaacacaccaaagacgacggcagtgtcagtgtcagtgtcgctgctgccgacttgccc  175
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116  aacgcatcgaacacaccaaagacgacggcagtgtcagtgtcagtgtcgctgctgccgacttgccc  175

Qy     176  aactggttcaaagcgaaggccggcagtgtcgtcaatattcaggcagccccgcccgcgca  235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     176  aactggttcaaagcgaaggccggcagtgtcgtcaatattcaggcagccccgcccgcgca  235

Qy     236  cccaaaacgcagcgcagcgcaatgccgaaaccgatccgaccgcttgccgacagcgaccgt  295
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     236  cccaaaacgcagcgcagcgcaatgccgaaaccgatccgaccgcttgccgacagcgaccgt  295

Qy     296  tctacgaattttcaaacgcctgtcccgaacatgcccgaaatccccagaagaagcag  355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     296  tctacgaattttcaaacgcctgtcccgaacatgcccgaaatccccagaagaagcag  355

Qy     356  atgacggcggattgaacttcggttcgggcttcatcatcagcaaaaacgctacatcctga  415
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     356  atgacggcggattgaacttcggttcgggcttcatcatcagcaaaaacgctacatcctga  415

Qy     416  ccaataccacgtcgttgccgtatggcagtatcaaagtcctgctcaacgacaagcgcg  475
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     416  ccaataccacgtcgttgccgtatggcagtatcaaagtcctgctcaacgacaagcgcg  475

Qy     476  aatataccgccaaactcatcggttcggatgtccaatccgatgtccgccctctgaaaatcg  535
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     476  aatataccgccaaactcatcggttcggatgtccaatccgatgtccgccctctgaaaatcg  535

Qy     536  acgcaacggaagagctcaccgtcgtcaatcggcaatcccaaaaatttgaaaccggggcg  595
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     536  acgcaacggaagagctcaccgtcgtcaatcggcaatcccaaaaatttgaaaccggggcg  595

Qy     596  aatgggtcgcgctgccatcggcgccgccctgcgttgacaacagcgtgaccgccggcatcg  655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     596  aatgggtcgcgctgccatcggcgccgccctgcgttgacaacagcgtgaccgccggcatcg  655

Qy     656  tgtccgccaaagcagaagctgccacgaacagctacacccttcatccaaccgacg  715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     656  tgtccgccaaagcagaagctgccacgaacagctacacccttcatccaaccgacg  715

Qy     716  ttgccatcaatccggcaatccggcgccgctgtccaacttaaaggacaggtcgtcg  775
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     716  ttgccatcaatccggcaatccggcgccgctgtccaacttaaaggacaggtcgtcg  775

Qy     776  gcatcaattcgcaaatatacagccggcggcggcgagattcatggcatcctcttgccatcc  835
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     776  gcatcaattcgcaaatatacagccggcggcggcgagattcatggcatcctcttgccatcc  835

Qy     836  cgattgacgttgcatgaagtgcgccgaacagctgaaaaacaccggcaaagtccaacgcg  895
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     836  cgattgacgttgcatgaagtgcgccgaacagctgaaaaacaccggcaaagtccaacgcg  895

Qy     896  gacaactgggcgtgattattcaggaagtatccacgtttggcacagtcgttcggtctcg  955
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     896  gacaactgggcgtgattattcaggaagtatccacgtttggcacagtcgttcggtctcg  955

Qy     956  ataaagccagcggcggcattgattgccaaaatccttccggcagccccgcagaacgtgccg  1015
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     956  ataaagccagcggcggcattgattgccaaaatccttccggcagccccgcagaacgtgccg  1015

Qy    1016  gcctgcagcggcggcggcgacatgtcctcagctcgacgcggagaaatacgttcttccgcg  1075
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1016  gcctgcagcggcggcggcgacatgtcctcagctcgacgcggagaaatacgttcttccgcg  1075

Qy    1076  accttccgtcatggtcggcgccattacgccggaaaagaagtcagcctcggcgtatggc  1135
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1076  accttccgtcatggtcggcgccattacgccggaaaagaagtcagcctcggcgtatggc  1135
```

```
Db    1264  gcagaacgcgcaggctcaaggcacggcgcgacgaatcctcagccgtcaggcgcaagtcccgt  1323
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    1393  caa  1395
            |||
Db    1324  caa  1326
```

RESULT   4

ID   Z54510 standard; DNA; 1500 BP.
XX
AC   Z54510;
XX
DT   21-MAR-2000 (first entry)
XX
DE   Neisseria gonorrhoeae ORF 986 partial DNA sequence SEQ ID NO:2967.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria gonorrhoeae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999; 99WO-US09346.
XX
PR   01-MAY-1998; 98US-0083758.
PR   31-JUL-1998; 98US-0094869.
PR   02-SEP-1998; 98US-0098994.
PR   02-SEP-1998; 98US-0099062.
PR   09-OCT-1998; 98US-0103749.
PR   09-OCT-1998; 98US-0103794.
PR   09-OCT-1998; 98US-0103796.
PR   25-FEB-1999; 99US-0121528.
XX
PA   (CHIR) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI   Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI   Tettelin H, Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; Y75748.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1389-1390; 1453pp; English.
XX
CC   Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC   novel Neisseria meningitis and N. gonorrheae polynucleotides and
CC   polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1500 BP; 384 A; 472 C; 379 G; 265 T; 0 other;

Query Match   89.5%; Score 1248; DB 21; Length 1500;
Best Local Similarity   99.9%; Pred. No. 0;

```
CC   chain reaction amplification. The NGSP polypeptide can also be used
CC   in screening assays to identify agents and compounds which useful as
CC   diagnostic, prophylactic or therapeutic agents against Neisseria
CC   infection.
XX
SQ   Sequence 1242 BP; 316 A; 400 C; 312 G; 214 T; 0 other;

Query Match             89.0%;  Score 1242;  DB 21;  Length 1242;
Best Local Similarity   100.0%;  Pred. NO. 0;
Matches 1242;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    154 atgctgctgccgacttgccaactgttcaaagcgaggccggcagtcgtcaatatt 213
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgctgctgccgacttgccaactgttcaaagcgaggccggcagtcgtcaatatt 60

Qy    214 caggcagcaccccgcccccgcgccccaaaacggcagcggcaatgccgaaaccgattccgac 273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 caggcagcaccccgcccccgcgccccaaaacggcagcggcaatgccgaaaccgattccgac 120

Qy    274 ccgcttgccgacgcgacccgttctacgaatttcaaagcgctcgtccgacatgccc 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccgcttgccgacgcgacccgttctacgaatttcaaagcgctcgtccgacatgccc 180

Qy    334 gaaatccccaagaagaagcagatgacgggattgaacttcgttcgggcttcatcatc 393
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gaaatccccaagaagaagcagatgacgggattgaacttcgttcgggcttcatcatc 240

Qy    394 agcaaaacggctacatcctgaccaatcacccacgtcgttgccgatgggcagtatcaaa 453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 agcaaaacggctacatcctgaccaatcacccacgtcgttgccgatgggcagtatcaaa 300

Qy    454 gtcctgctcaacgacaagcggcgaatatacacgccaaactcatcgttcgatgtccaatcc 513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gtcctgctcaacgacaagcggcgaatatacacgccaaactcatcgttcgatgtccaatcc 360

Qy    514 gatgtcgccctctgaaaatcgacgcaacggaagagctcacccgtctgtcaaaatcggcaat 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gatgtcgccctctgaaaatcgacgcaacggaagagctcacccgtctgtcaaaatcggcaat 420

Qy    574 cccaaaaatttgaaaccggccgaatgggtcgctcgccatccggccgcgccccttcgcttgac 633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cccaaaaatttgaaaccggccgaatgggtcgctcgccatccggccgcgccccttcgcttgac 480

Qy    634 aacagcgtgaccgccggacatcgtgtccgcaaggcagcagctgccaagaaagctac 693
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aacagcgtgaccgccggacatcgtgtccgcaaggcagcagctgccaagaaagctac 540

Qy    694 acacccttcatccaaccgacgttgccatcaatccggccacatccggccgcccctgttc 753
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 acacccttcatccaaccgacgttgccatcaatccggccacatccggccgcccctgttc 600

Qy    754 aacttaaaaggacaggtcgtcggcacaatccaacatcaattcgcaaatatacagcgcggcggatcc 813
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aacttaaaaggacaggtcgtcggcacaatccaacatcaattcgcaaatatacagcgcggcggatcc 660

Qy    814 atgggcatctccttgccatcccgattgacgtgccatgaatgtcgcgaacagctgaaa 873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 atgggcatctccttgccatcccgattgacgtgccatgaatgtcgcgaacagctgaaa 720

Qy    874 aacaccggcaaagtccaacgcggacaactgggcgtgattattccaggaagtatcctacggt 933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 aacaccggcaaagtccaacgcggacaactgggcgtgattattccaggaagtatcctacggt 780

Qy    934 ttggcacagtcgttcggtctgataaagccggccattgattggccaaaatccttccc 993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ttggcacagtcgttcggtctgataaagccggccattgattggccaaaatccttccc 840

Qy    994 ggcagccccagaacgtgccgcctgcagggccgacatcgtctctcagcctcgacggc 1053
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ggcagccccagaacgtgccgcctgcagggccgacatcgtctctcagcctcgacggc 900

Qy   1054 ggagaaatacgttcttccgcgcacctcccgtcatgtcggcgccattaccgccggaaaa 1113

Qy   1136 gcaaaggcggagaaatcacacatcaaagccaagctggcaacgccgccgagcatacccgcg 1195
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   1136 gcaaaggcggagaaatcacacatcaaagccaagctggccacgccgccgagcatacccgcg 1195

Qy   1196 catcatccaaaacagatgaagcccctacaccgaacagcaatccggtacgttctcggtcg 1255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1196 catcatccaaaacagatgaagcccctacaccgaacagcaatccggtacgttctcggtcg 1255

Qy   1256 aatccgcaggcattaccctcagacacatccgacacagcagcggcaaacacctcgtcgtcg 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1256 aatccgcaggcattaccctcagacacatccgacacagcagcggcaaacacctcgtcgtcg 1315

Qy   1316 tacgggtttccgacgcggcagcagaacgcgcaggcttaaggc 1354
          ||||||||||||||||||||||||||||||||||||||||||
Db   1316 tacgggtttccgacgcggcagcagaacgcgcaggcttaaggc 1354


RESULT 5
Z93415

ID   Z93415 standard; DNA; 1242 BP.
XX
AC   Z93415;
XX
DT   24-JUL-2000 (first entry)
XX
DE   NGSP polypeptide coding sequence.
XX
KW   NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW   cellular matrix; adherence domain; ligand; detection; diagnosis;
KW   screening; probe; primer; prophylaxis; therapy; ss.
XX
OS   Neisseria gonorrhoeae.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1242
FT                  /tag= a
FT                  /product= NGSP polypeptide
XX
PN   WO200012133-A1.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US20070.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ, Harris AM;
XX
DR   WPI; 2000-237782/20.
DR   P-PSDB; Y83151.
XX
PT   Non-cytosolic NGSP polypeptide and polynucleotide sequence from
PT   Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT   infections
XX
PS   Claim 38; Page 62-63; 68pp; English.
XX
CC   The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC   Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC   domains for extracellular matrix proteins. Using the NGSP polypeptide
CC   as a vaccine produces antibodies which inhibit binding of N.
CC   gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC   subsequent invasion. The NGSP polypeptide and its peptide fragments
CC   can be used to immunise an animal and produce an immune response.
CC   They can also be used as ligands to detect antibodies elicited in
CC   response to Neisseria infections and also as antigens or immunogens
CC   for inducing Neisseria-specific antibodies which are useful in
CC   immunoassays to detect Neisseria in biological specimens. Nucleotides
CC   encoding NGSP or its fragments can be used as probes to identify
CC   Neisseria in biological specimens by hybridization or polymerase
```

```
Db   901 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         ggagaaatacgttctccggacgaccttcccgtcatggtcggcgccattacgcaccggaaaa 960

Qy  1114 gagtcagctcggcgtatggcgcaaagcggagaaatcacaatcaaagccaagctggc 1173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 gagtcagctcggcgtatggcgcaaagcggagaaatcacaatcaaagccaagctgggc 1020

Qy  1174 aacgcgccggcagcataccggcgcatcatccaaaacagatgagcccctaccaccgaacag 1233
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 aacgcgccggcagcataccggcgcatcatccaaaacagatgagcccctaccaccgaacag 1080

Qy  1234 caatccgtacgttctcggtcgaatccgcaggcattaccccttcagacacataccgacagc 1293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 caatccgtacgttctcggtcgaatccgcaggcattaccccttcagacacataccgacagc 1140

Qy  1294 agcggcaaacacctcgtcgtctacgggtttccgacgcggcagaacgcgcaggcttaagg 1353
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 agcggcaaacacctcgtcgtctacgggtttccgacgcggcagaacgcgcaggcttaagg 1200

Qy  1354 cacggggcgacgaaatcctagccgtcaggcaagtccccgtcaa 1395
         ||||||||||||||||||||||||||||||||||||||||||
Db  1201 cacggggcgacgaaatcctagccgtcaggcaagtccccgtcaa 1242
```

```
RESULT
Z51533
ID    Z51533 standard; DNA; 1347 BP.
XX
AC    Z51533;
XX
DT    03-JUL-2000  (first entry)
XX
DE    Neisseria meningitidis NMASP protein-1 encoding DNA.
XX
KW    NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW    anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW    Neisserial infection; meningitidis; septicaemia; ds.
XX
OS    Neisseria meningitidis.
XX
FH    Key              Location/Qualifiers
FT    CDS              1..1347
FT                     /tag= a
FT                     /product= "Neisseria meningitidis protein"
FT                     /transl_except= (pos:499..501, aa: Xaa)
FT                     /note= "Xaa is unknown"
XX
PN    WO200012535-A2.
XX
PD    09-MAR-2000.
XX
PF    01-SEP-1999; 99WO-US19663.
XX
PR    01-SEP-1998; 98US-0098685.
XX
PA    (ANTE-) ANTEX BIOLOGICS INC.
XX
PI    Jackson WJ, Harris AM;
XX
DR    WPI; 2000-256581/22.
DR    P-PSDB; Y70409.
XX
PT    Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT    antibodies, useful in vaccines against infection -
XX
PS    Claim 41; Page 63; 75pp; English.
XX
CC    The present sequence is a DNA encoding Neisseria meningitidis NMASP
CC    protein. NMASP is a non-cytosolic protein, with antibacterial and
CC    antiinflammatory activity. It shows sequence similarity to E. coli
CC    DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC    antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC    anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
```

```
CC    may be used for diagnosis, therapy or prophylaxis of Neisserial
CC    infections such as, bacterial meningitidis and septicaemia.
CC    Note: The protein sequence represented in SEQ ID NO:2 of the
XX    specification is erroneous"
SQ    Sequence 1347 BP; 338 A; 428 C; 344 G; 236 T; 1 other;

Query Match          12.5%;  Score 175; DB 21; Length 1347;
Best Local Similarity   99.6%;  Pred. No. 1e-80;
Matches  225;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   174 ccaactggttcaaagcaaggcccggcagtcgtcaatattcaggcagcccgcccccgcg 233
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 ccaactggttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccgcccccgcg 80

Qy   234 cacccaaaacggcagcggcaatgccgaaccgattccgacccgcttgccgacaggcgacc 293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 cacccaaaacggcagcagcaatgccgaaaccgattccgacccgcttgccgacagcgacc 140

Qy   294 gttctacgaattttcaaacgcctgtcccgaacatgcccgaaatcccccaagaagaagc 353
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 gttctacgaattttcaaacgcctgtcccgaacatgcccgaaatcccccaagaagaagc 200

Qy   354 agatgacggcggattgaacttcggttcggcttcatcatcagcaa 399
         ||||||||||||||||||||||||||||||||||||||||||
Db   201 agatgacggcggattgaacttcggttcgggcttcatcatcagcaa 246
```

```
RESULT
Z33305   7
ID    Z33305 standard; DNA; 1500 BP.
XX
AC    Z33305;
XX
DT    21-FEB-2000  (first entry)
XX
DE    Neisseria meningitidis strain ATCC 13090 BASB013 nucleotide sequence.
XX
KW    Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW    antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW    invasive bacterial disease; antibacterial; ss.
XX
OS    Neisseria meningitidis.
XX
PN    WO955872-A1.
XX
PD    04-NOV-1999.
XX
PF    20-APR-1999;  99WO-EP02765.
XX
PR    23-APR-1998;  98GB-0008734.
XX
PA    (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI    Ruelle J;
XX
DR    WPI; 2000-052809/04.
DR    P-PSDB; Y52993.
XX
PT    Novel polynucleotides and polypeptides from Neisseria meningitis used
PT    to prepare vaccines against bacterial infections -
XX
PS    Claim 15; Page 73-74; 94pp; English.
XX
CC    The present sequence encodes a BASB013 polypeptide isolated from
CC    Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC    employed as research reagents and material for the discovery of
CC    treatments and diagnostics for diseases, particularly human diseases.
CC    They can be used for diagnosis of disease, staging of disease, or
CC    determining response of an infectious organism to drugs. The
CC    polynucleotides may be used as a source for hybridisation probes, and
CC    for screening of genetic mutations, serotype, organism or strain
```

CC　identification, identification of mutation in BASB013 sequences, and as
CC　components of arrays which are useful for diagnostic and prognostic
CC　purposes. The polypeptides can be used to produce antibodies. The
CC　polypeptides can also be used in vaccine formulations, and to identify
CC　agonists and antagonists. The polypeptides, antibodies, agonists and
CC　antagonists (which are bacteristatic) are used for the treatment and
CC　prevention of diseases such as upper respiratory tract infection,
CC　invasive bacterial diseases such as bacteraemia and meningitis, and for
CC　the development and screening of antibacterial drugs. They are also used
CC　in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC　on in-dwelling devices, or to extracellular proteins on wounds, and to
CC　thus prevent tissue damage and/or block the normal progression of
CC　pathogenesis in infections initiated other than by the implantation of
CC　in-dwelling devices or by other surgical techniques.
XX
SQ　Sequence 1500 BP; 370 A; 468 C; 393 G; 269 T; 0 other;

Query Match　　　　　12.5%;　　Score 175;　DB 21;　Length 1500;
Best Local Similarity　99.6%;　Pred. No. 1e-80;
Matches　225;　Conservative　0;　Mismatches　1;　Indels　0;　Gaps　0;

Qy　174　ccaactggttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccccgccgcg　233
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　174　ccaactggttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccccgccgcg　233

Qy　234　cacccaaaacggcagcggcagcggcaatgccgaaaccgattccgaccgcttgccgacagcgaccc　293
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　234　cacccaaaacggcagcggcagcggcaatgccgaaaccgattccgaccgcttgccgacagcgaccc　293

Qy　294　gttctacgaattttcaaacgcctcgtcccgaacatgcccgaaatgcccgaagaagagc　353
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||
Db　294　gttctacgaattttcaaacgcctcgtcccgaatgcccgaaatgcccgaagaagagc　353

Qy　354　agatgacggcggattgaacttcggttcgggcttcatcatcagcaaa　399
　　　　　||||||||||||||||||||||||||||||||||||||||||||
Db　354　agatgacggcggattgaacttcggttcgggcttcatcatcagcaaa　399

RESULT　8
Z33306
ID　Z33306 standard; DNA; 1500 BP.
XX
AC　Z33306;
XX
DT　21-FEB-2000　(first entry)
XX
DE　Neisseria meningitidis strain ATCC 13090 BASB013 nucleotide sequence.
XX
KW　Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW　antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW　invasive bacterial disease; antibacterial; ss.
XX
OS　Neisseria meningitidis.
XX
PN　WO955872-A1.
XX
PD　04-NOV-1999.
XX
PF　20-APR-1999;　99WO-EP02765.
XX
PR　23-APR-1998;　98GB-0008734.
XX
PA　(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI　Ruelle J;
XX
DR　WPI; 2000-052809/04.
DR　P-PSDB; Y52994.
XX
PT　Novel polynucleotides and polypeptides from Neisseria meningitis used
PT　to prepare vaccines against bacterial infections -
XX

PS　Claim 12; Page 76; 94pp; English.
XX
CC　The present sequence encodes a BASB013 polypeptide isolated from
CC　Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC　employed as research reagents and material for the discovery of
CC　treatments and diagnostics for diseases, particularly human diseases.
CC　They can be used for diagnosis of disease, staging of disease, or
CC　determining response of an infectious organism to drugs. The
CC　polynucleotides may be used as a source for hybridisation probes, and
CC　for screening of genetic mutations, serotype, organism or strain
CC　identification, identification of mutation in BASB013 sequences, and as
CC　components of arrays which are useful for diagnostic and prognostic
CC　purposes. The polypeptides can be used to produce antibodies. The
CC　polypeptides can also be used in vaccine formulations, and to identify
CC　agonists and antagonists. The polypeptides, antibodies, agonists and
CC　antagonists (which are bacteristatic) are used for the treatment and
CC　prevention of diseases such as upper respiratory tract infection,
CC　invasive bacterial diseases such as bacteraemia and meningitis, and for
CC　the development and screening of antibacterial drugs. They are also used
CC　in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC　on in-dwelling devices, or to extracellular proteins on wounds, and to
CC　thus prevent tissue damage and/or block the normal progression of
CC　pathogenesis in infections initiated other than by the implantation of
CC　in-dwelling devices or by other surgical techniques.
XX
SQ　Sequence 1500 BP; 370 A; 469 C; 392 G; 269 T; 0 other;

Query Match　　　　　12.5%;　　Score 175;　DB 21;　Length 1500;
Best Local Similarity　99.6%;　Pred. No. 1e-80;
Matches　225;　Conservative　0;　Mismatches　1;　Indels　0;　Gaps　0;

Qy　174　ccaactggttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccccgccgcg　233
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　174　ccaactggttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccccgccgcg　233

Qy　234　cacccaaaacggcagcggcagcggcaatgccgaaaccgattccgaccgcttgccgacagcgaccc　293
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　234　cacccaaaacggcagcggcagcggcaatgccgaaaccgattccgaccgcttgccgacagcgaccc　293

Qy　294　gttctacgaattttcaaacgcctcgtcccgaacatgcccgaaatgcccgaagaagagc　353
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||
Db　294　gttctacgaattttcaaacgcctcgtcccgaatatgcccgaaatgcccgaagaagagc　353

Qy　354　agatgacggcggattgaacttcggttcgggcttcatcatcagcaaa　399
　　　　　||||||||||||||||||||||||||||||||||||||||||||
Db　354　agatgacggcggattgaacttcggttcgggcttcatcatcagcaaa　399

RESULT　9
254511
ID　254511 standard; DNA; 1500 BP.
XX
AC　254511;
XX
DT　21-MAR-2000　(first entry)
XX
DE　Neisseria meningitidis ORF 986 partial DNA sequence SEQ ID NO:2969.
XX
KW　Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW　antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW　antibacterial; gene therapy; ds.
XX
OS　Neisseria meningitidis.
XX
PN　WO957280-A2.
XX
PD　11-NOV-1999.
XX
PF　30-APR-1999;　99WO-US09346.
XX
PR　01-MAY-1998;　98US-0083758.
PR　31-JUL-1998;　98US-0094869.

PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI   Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI   Tettelin H, Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; Y75749.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1390; 1453pp; English.
XX
CC   253015 to 253536, 254536, 254577 to 254615, and Y74253 to Y75941 represent
CC   novel Neisseria meningitidis and N.gonorrheae polynucleotides and
CC   polypeptides. 254537 to 254576 and 254616 to 255473 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1500 BP; 386 A; 469 C; 382 G; 263 T; 0 other;

Query Match   11.2%;   Score 156; DB 21;   Length 1500;
Best Local Similarity   99.0%;   Pred. No. 6.9e-71;
Matches   306;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

Qy   991 cccggcagcccccgagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgac 1050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   991 cccggcagcccccgagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgac 1050

Qy   1051 ggcgggaaatacgttcttccggcgacctccgtcatgtcggcgccattacgccggga 1110
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1051 ggcgggaaatacgttcttccggcgacctccgttatgtcggcgccattacgccggga 1110

Qy   1111 aagaagtcagcctcggcgtatggcgcaaggcgaagaaatcacaatcaaagtcaagctg 1170
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1111 aagaagtcagcctcggcgtatggcgcaaggcgaagaaatcacaatcaaagtcaagctg 1170

Qy   1171 ggcaacgccgccgagcataccggcgcatcatccaaaacagatgaagccctacaccgaa 1230
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1171 ggcaacgccgccgagcatatccggtcgaatccgcaggcattaccccttcagacacataccgac 1230

Qy   1231 cagcaatccggtacgttctcggtcgaatccgcaggcattaccccttcagacacataccgac 1290
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1231 cagcaatccggtacgttctcggtcgaatccgcaggcattaccccttcagacacataccgac 1290

Qy   1291 agcagccgc 1299
          |||||||||
Db   1291 agcagccgc 1299

RESULT   10
233307
ID   233307 standard; DNA; 1500 BP.
XX
AC   233307;

XX
DT   21-FEB-2000 (first entry)
XX
DE   Neisseria meningitidis strain H44/76 BASB013 nucleotide sequence.
XX
KW   Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW   antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW   invasive bacterial disease; antibacterial; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO0955872-A1.
XX
PD   04-NOV-1999.
XX
PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
XX
PA   (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-C52809/04.
DR   P-PSDB; Y52995.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections -
XX
PS   Claim 12; Page 78-79; 94pp; English.
XX
CC   The present sequence encodes a BASB013 polypeptide isolated from
CC   Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC   employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence 1500 BP; 385 A; 471 C; 382 G; 262 T; 0 other;

Query Match   11.2%;   Score 156; DB 21;   Length 1500;
Best Local Similarity   99.0%;   Pred. No. 6.9e-71;
Matches   306;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

Qy   991 cccggcagcccccgagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgac 1050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   991 cccggcagcccccgagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgac 1050

Qy   1051 ggcgggaaatacgttcttccggcgacctccgtcatgtcggcgccattacgccggga 1110
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1051 ggcgggaaatacgttcttccggcgacctccgttatgtcggcgccattacgccggga 1110

Qy   1111 aagaagtcagcctcggcgtatggcgcaaggcgaagaaatcacaatcaaagtcaagctg 1170
          |||||||||||||||||||||||| ||||||||||||||||||||||||| |||||||
Db   1111 aagaagtcagcctcggcgtatggcgcaaggcgaagaaatcacaatcaaagtcaagctg 1170

```
Qy   1171  ggaacgccgccggcgcagcataccggcatcatccaaaacagatgaagccccctacaccgaa  1230
           ||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db   1171  ggaacgccgccggcgcagcatccggcatcatccaaaacagatgaagccccctacaccgaa  1230

Qy   1231  cagcaatccggtacgttccggtcgaatccgcaggcattaccctcagacacataccggac  1290
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1231  cagcaatccggtacgttccggtcgaatccgcaggcattaccctcagacacataccggac  1290

Qy   1291  agcagcggc  1299
           |||||||||
Db   1291  agcagcggc  1299

RESULT  11
Z93416
ID   Z93416 standard; DNA; 153 BP.
XX
AC   Z93416;
XX
DT   24-JUL-2000  (first entry)
XX
DE   NGSP coding sequence fragment.
XX
KW   NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW   cellular matrix; adherence domain; ligand; detection; diagnosis;
KW   screening; probe; primer; prophylaxis; therapy; ss.
XX
OS   Neisseria gonorrhoeae.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1242
FT                  /tag= a
FT                  /product= NGSP polypeptide
XX
PN   WO20001213-A1.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;  99WO-US20070.
XX
PR   01-SEP-1998;  98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-237782/20.
DR   P-PSDB; Y83152.
XX
PT   Non-cytosolic NGSP polypeptide and polynucleotide sequence from
PT   Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT   infections
XX
PS   Claim 38; Page 65; 68pp; English.
XX
CC   The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC   Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC   domains for extracellular matrix proteins. Using the NGSP polypeptide
CC   as a vaccine produces antibodies which inhibit binding of N.
CC   gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC   subsequent invasion. The NGSP polypeptide and its peptide fragments
CC   can be used to immunise an animal and produce an immune response.
CC   They can also be used as ligands to detect antibodies elicited in
CC   response to Neisseria infections and also as antigens or immunogens
CC   for inducing Neisseria specific antibodies which are useful in
CC   immunoassays to detect Neisseria in biological specimens. Nucleotides
CC   encoding NGSP or its fragments can be used as probes to identify
CC   Neisseria in biological specimens by hybridization or polymerase
CC   chain reaction amplification. The NGSP polypeptide can also be used
CC   in screening assays to identify agents and compounds which useful as
CC   diagnostic, prophylactic or therapeutic agents against Neisseria
CC   infection.
XX
```

```
SQ   Sequence 153 BP; 43 A; 38 C; 41 G; 31 T; 0 other;

     Query Match             11.0%;  Score 153;  DB 21;  Length 153;
     Best Local Similarity  100.0%;  Pred. No. 2.3e-69;
     Matches 153;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1  gtgttcaaaaataccaatactccgctttggccgcactgtgtgccgcccttgctggccaggc   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  gtgttcaaaaataccaatactccgctttggccgcactgtgtgccgcccttgctggccaggc   60

Qy    61  tgcgaaaggcaggcagcgctttttcgtgcggacaaaaagaagaagcatccttcgtagaacgc  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  tgcgaaaggcaggcagcgctttttcgtgcggacaaaaagaagaagcatccttcgtagaacgc  120

Qy   121  atcgaacaccaaagacgacggcagtgtcagt  153
          |||||||||||||||||||||||||||||||
Db   121  atcgaacaccaaagacgacggcagtgtcagt  153

RESULT  12
Z54512
ID   Z54512 standard; DNA; 1500 BP.
XX
AC   Z54512;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 986 partial DNA sequence SEQ ID NO:2971.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;  99WO-US09346.
XX
PR   01-MAY-1998;  98US-0083758.
PR   31-JUL-1998;  98US-0094869.
PR   02-SEP-1998;  98US-0098994.
PR   02-SEP-1998;  98US-0099062.
PR   09-OCT-1998;  98US-0103749.
PR   09-OCT-1998;  98US-0103794.
PR   09-OCT-1998;  98US-0103796.
PR   25-FEB-1999;  99US-0121528.
XX
PA   (CHIR-) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V, Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E, Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; Y75750.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1392; 1453pp; English.
XX
CC   253015 to 254536, 254577 to 254615, and Y74253 to Y75941 represent
CC   novel Neisseria meningitis and N. gonorrhoeae polynucleotides and
CC   polypeptides. 254537 to 254576 and 254616 to 254673 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
XX
```

CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1500 BP; 382 A; 469 C; 382 G; 265 T; 2 other;

Query Match   9.9%;   Score 138; DB 21;   Length 1500;
Best Local Similarity   99.5%;   Pred. No. 1.4e-61;
Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   174 ccaactggttcaaagcgaaggccggagtcgtcaatattcaggcagcccccgccg 233
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   174 ccaactggttcaaagcgaaggcccggagtcgtcaatattcaggcagcccccgccg 233

Qy   234 cacccaaaacggcacgcagcgcaatgccgaaaccgattccgacccgcttgccgacagcgaccc 293
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   234 cacccaaaacggcacgcagcgcaatgccgaaaccgattccgacccgcttgccgacagcgaccc 293

Qy   294 gttctacgaattttcaaaacgcctcgtccgaacatgcccgaatatcccccaagaagaagc 353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   294 gttctacgaattttcaaaacgcctcgtccgaacatgcccgaatatcccccaagaagaagc 353

Qy   354 agatgacgg 362
          |||||||||
Db   354 agatgacgg 362

CC   The present sequence encodes a conserved BASB013-C polypeptide isolated
CC   from Neisseria meningitidis. BASB013 polynucleotides and polypeptides
CC   may be employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The

CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as bacteraemia and meningitis, and for
CC   invasive bacterial diseases such as upper respiratory tract infection,
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence 1110 BP; 280 A; 350 C; 279 G; 201 T; 0 other;

Query Match   8.1%;   Score 113; DB 21;   Length 1110;
Best Local Similarity 100.0%;   Pred. No. 1.1e-48;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   439 atgggcagtatcaaagtcctgtcaacgacaacgcgcgaatataccgccaaactcatcggt 498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   439 atgggcagtatcaaagtcctgtcaacgacaacgcgcgaatataccgccaaactcatcggt 498

Qy   499 tcggatgtccaatccgatgtcgccttctgaaaatcgacgcaacggaagagct 551
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   499 tcggatgtccaatccgatgtcgccttctgaaaatcgacgcaacggaagagct 551

CC   The present sequence encodes a variable BASB013-V polypeptide isolated
CC   from Neisseria meningitidis. BASB013 polynucleotides and polypeptides
CC   may be employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.

CC　They can be used for diagnosis of disease, staging of disease, or
CC　determining response of an infectious organism to drugs. The
CC　polynucleotides may be used as a source for hybridisation probes, and
CC　for screening of genetic mutations, serotype, organism or strain
CC　identification, identification of mutation in BASB013 sequences, and as
CC　components of arrays which are useful for diagnostic and prognostic
CC　purposes. The polypeptides can be used to produce antibodies. The
CC　polypeptides can also be used in vaccine formulations, and to identify
CC　agonists and antagonists. The polypeptides, antibodies, agonists and
CC　antagonists (which are bacteristatic) are used for the treatment and
CC　prevention of diseases such as upper respiratory tract infection,
CC　invasive bacterial diseases such as bacteraemia and meningitis, and for
CC　the development and screening of antibacterial drugs. They are also used
CC　in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC　on in-dwelling devices, or to extracellular proteins on wounds, and to
CC　thus prevent tissue damage and/or block the normal progression of
CC　pathogenesis in infections initiated other than by the implantation of
CC　in-dwelling devices or by other surgical techniques.
XX
SQ　Sequence 390 BP; 106 A; 120 C; 102 G; 62 T; 0 other;

Query Match　　　　7.8%;　Score 109;　DB 21;　Length 390;
Best Local Similarity　100.0%;　Pred. No. 1.3e-46;
Matches 109;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　1191 cggcgcatccatccaaaacagatgaagcccctaccggacacagcaatccggtacgttctc 1250
　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　81 cggcgcatccatccaaaacagatgaagcccctaccggacacagcaatccggtacgttctc 140

Qy　1251 ggtcgaatccgcaggcattaccccttcagacacataccgacgagcggc 1299
　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||
Db　　141 ggtcgaatccgcaggcattaccccttcagacacataccgacgagcggc 189

RESULT 15
251542
ID　251542 standard; DNA; 153 BP.
XX
AC　251542;
XX
DT　03-JUL-2000　(first entry)
XX
DE　Neisseria meningitidis NMASP protein N-terminal fragment encoding DNA.
XX
KW　NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW　anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW　Neisserial infection; meningitidis; septicaemia; ds.
XX
OS　Neisseria meningitidis.
XX
PN　WO200012535-A2.
XX
PD　09-MAR-2000.
XX
PF　01-SEP-1999;　99WO-US19663.
XX
PR　01-SEP-1998;　98US-0098685.
XX
PA　(ANTE-) ANTEX BIOLOGICS INC.
XX
PI　Jackson WJ, Harris AM;
XX
DR　WPI; 2000-255581/22.
DR　P-PSDB; Y70415.
XX
PT　Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT　antibodies, useful in vaccines against infection -
XX
PS　Claim 42; Page 74; 75pp; English.
XX
CC　The present sequence is a DNA encoding Neisseria meningitidis NMASP
CC　protein N-terminal fragment. NMASP is a non-cytosolic protein, with

CC　antibacterial and antiinflammatory activity. It shows sequence similarity
CC　to E. coli DegP (HtrA) protein. NMASP proteins can be used as ligands to
CC　detect antibodies elicited in response to N. meningitidis infections.
CC　Cytotoxic anti-NMASP antibodies can be used as vaccines. NMASP proteins
CC　and DNA may be used for diagnosis, therapy or prophylaxis of Neisserial
CC　infections such as, bacterial meningitidis and septicaemia.
XX
SQ　Sequence 153 BP; 44 A; 42 C; 38 G; 29 T; 0 other;

Query Match　　　　2.7%;　Score 38;　DB 21;　Length 153;
Best Local Similarity　100.0%;　Pred. No. 6.4e-10;
Matches　38;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　73 ggcagcttttcggtgcggacaaaaagaagaagcatcctt 110
　　　　 |||||||||||||||||||||||||||||||||||||||
Db　　73 ggcagcttttcggtgcggacaaaaagaagaagcatcctt 110

Search completed: March 22, 2001, 02:52:06
Job time: 4456 sec

us-09-388-090-3.oligo25.rng

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 01:36:30 ; Search time 58.8 Seconds
                                            (without alignments)
                                            3823.446 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaataccaata.........tcaggcaagtcccgtcaa 1395

Scoring table: OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0

Searched: 280836 seqs, 80580151 residues

Word size :   25

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   Issued_Patents_NA:*
                 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                 3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
                 4: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                 5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

           Pred. No. is the number of results predicted by chance to have a
           score greater than or equal to the score of the result being printed,
           and is derived by analysis of the total score distribution..

                                  SUMMARIES

                          %
Result   Query
No.   Score Match Length DB ID   Description

No matches found

Search completed: March 22, 2001, 02:50:12
Job time: 4422 sec

us-09-388-090-3.oligo25.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 01:22:25 ; Search time 885.72 Seconds
                                                (without alignments)
                                                11036.722 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaataccaata.........tcaggcaagtcccgtcaa 1395

Scoring table: OLIGO_NUC
                     Gapop 60.0 , Gapext 60.0

Searched: 7991742 seqs, 3503743858 residues

Word size :   25

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   EST:*

  1: gb_est1:*   44: em_esthum2:*
  2: gb_est2:*   45: em_esthum3:*
  3: gb_est3:*   46: em_esthum4:*
  4: gb_est4:*   47: em_esthum5:*
  5: gb_est5:*   48: em_esthum6:*
  6: gb_est6:*   49: em_esthum7:*
  7: gb_est7:*   50: em_esthum8:*
  8: gb_est8:*   51: em_esthum9:*
  9: gb_est9:*   52: em_esthum10:*
 10: gb_est10:*   53: em_esthum11:*
 11: gb_est11:*   54: em_esthum12:*
 12: gb_est12:*   55: em_esthum13:*
 13: gb_est13:*   56: em_esthum14:*
 14: gb_est14:*   57: em_esthum15:*
 15: gb_est15:*   58: em_esthum16:*
 16: gb_est16:*   59: em_esthum17:*
 17: gb_est17:*   60: em_esthum18:*
 18: gb_est18:*   61: em_esthum19:*
 19: gb_est19:*   62: em_esthum20:*
 20: gb_est20:*   63: em_estin1:*
 21: gb_est21:*   64: em_estin2:*
 22: gb_est22:*   65: em_estin3:*
 23: gb_est23:*   66: em_estin4:*
 24: gb_est24:*   67: em_estov1:*
 25: gb_est25:*   68: em_estov2:*
 26: gb_est26:*   69: em_estpl1:*
 27: gb_est27:*   70: em_estpl2:*
 28: gb_est28:*   71: em_estpl3:*
 29: gb_est29:*   72: em_estpl4:*
 30: gb_est30:*   73: em_estpl5:*
 31: gb_est31:*   74: em_estro1:*
 32: gb_est32:*   75: em_estro2:*
 33: gb_est33:*   76: em_estro3:*
 34: gb_est34:*   77: em_estro4:*
 35: gb_est35:*   78: em_estro5:*
 36: gb_est36:*   79: em_estro6:*
 37: gb_est37:*   80: em_estro7:*
 38: gb_est38:*   81: em_estro8:*
 39: gb_est39:*   82: em_estro9:*
 40: gb_est40:*   83: em_estro10:*
 41: em_estba:*   84: em_estro11:*
 42: em_estfun:*   85: em_estro12:*
 43: em_esthum1:*   86: em_estro13:*
                                             87: gb_est41:*
                                             88: gb_est42:*
                                             89: gb_est43:*
                                             90: gb_est44:*
                                             91: gb_est45:*
                                             92: gb_est46:*
                                             93: gb_est47:*
                                             94: gb_est48:*
                                             95: gb_est49:*
                                             96: gb_est50:*
                                             97: gb_est51:*
                                             98: gb_est52:*
                                             99: gb_est53:*
                                            100: gb_est54:*
                                            101: gb_est55:*
                                            102: gb_est56:*
                                            103: gb_est57:*
                                            104: gb_est67:*
                                            105: gb_est68:*
                                            106: gb_est69:*
                                            107: gb_est70:*
                                            108: gb_est71:*
                                            109: gb_est72:*
                                            110: gb_est73:*
                                            111: gb_est74:*
                                            112: em_esthum21:*
                                            113: em_esthum22:*
                                            114: em_esthum23:*
                                            115: em_estom1:*
                                            116: em_estom2:*

```
117: em_estp16:*
118: em_estp17:*
119: em_estp18:*
120: em_estro14:*
121: em_estro15:*
122: em_estro16:*
123: em_estro17:*
124: em_estro18:*
125: em_estro19:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est75:*
136: gb_est76:*
137: gb_est77:*
138: gb_est78:*
139: gb_est79:*
140: gb_est80:*
141: gb_est81:*
142: gb_est82:*
143: gb_est83:*
144: gb_est84:*
145: gb_est85:*
146: gb_est86:*
147: gb_est87:*
148: gb_gss1:*
149: gb_gss2:*
150: gb_gss3:*
151: gb_gss4:*
152: em_gss1:*
153: em_gss2:*
154: em_gss3:*
155: em_gss4:*
156: gb_gss5:*
157: gb_gss6:*
158: gb_gss7:*
159: gb_gss8:*
160: gb_gss9:*
161: gb_gss10:*
162: gb_gss11:*
163: gb_gss12:*
164: gb_gss13:*
165: gb_gss14:*
166: gb_gss15:*
167: gb_gss16:*
168: gb_gss17:*
169: gb_gss18:*
170: gb_gss19:*
171: gb_gss20:*
172: gb_gss21:*
173: gb_gss22:*
174: gb_gss23:*
175: gb_gss24:*
176: em_gss5:*
177: em_gss6:*
178: em_gss7:*
179: em_gss8:*
180: em_gss9:*
181: em_gss10:*
182: em_gss11:*
183: em_gss12:*
184: em_gss13:*
185: em_gss14:*
186: em_gss15:*
187: em_gss16:*
188: em_gss17:*
189: em_gss18:*
```

```
190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                    SUMMARIES

                                         %
Result   Query
No.   Score   Match Length DB ID   Description
------ -----   ----- ------ --- --------------------------------------

No matches found

Search completed: March 22, 2001, 02:20:30
Job time: 3485 sec

us-09-388-090-3.oligo25.rst

Fri Mar 23 07:26:21 2001

```
OM of: US-09-388-090-3 to: A_Geneseq_36:*    out_format : pfs

Date: Mar 22, 2001 12:04 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool/US09388090/runat_20032001_092950_25784/app_query.fasta_1.1474
-DB=A_Geneseq_36 -QFMT=fastan -SUFFIX=oligo7.rag -GAPOP=4.500
-GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR_SCORE=quality
-THR_MIN=7 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09388090 -@CGN1_1_36 -NCPU=6 -ICPU=3
-LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-388-090-3
Query length: 1395
Database: A_Geneseq_36:*
Database sequences: 268485
Database length: 34193795
Search time (sec): 33.030000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:
Sequence                                              Strd Orig   ZScore   EScore   Len  ! Documentation
```

Sequence	Strd	Orig	ZScore	EScore	Len	!
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y83150	+	465.00	8134.18	0.0	465	!
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75748	+	451.00	7887.98	0.0	499	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y83151	+	414.00	7239.91	0.0	414	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y52996	+	127.00	2203.72	1.1e-115	370	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75749	+	127.00	2201.64	1.1e-115	499	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y52994	+	127.00	2201.64	1.1e-115	499	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y52995	+	127.00	2201.64	1.1e-115	499	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y52993	+	115.00	1991.04	5.8e-104	499	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y70409	+	103.00	1781.18	3.2e-92	448	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y70414	+	103.00	1780.77	3.1e-92	475	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y70413	+	103.00	1780.45	3.1e-92	498	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75750	+	103.00	1780.43	3.1e-92	499	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y83152	+	51.00	883.64	2.7e-41	51	! NC
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y52997	+	36.00	613.94	1.1e-26	129	! N
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y70415	+	19.00	322.02	5.2e-10	51	! Ne
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y29294	+	18.00	289.22	3.9e-09	460	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y83154	+	11.00	192.26	0.0410	11	! C
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:W60988	+	11.00	171.14	0.0293	231	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W35634	+	11.00	168.94	0.0283	317	! C
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:W81772	+	11.00	167.38	0.0276	397	! S
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W37309	+	11.00	165.90	0.0270	491	! F
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y83153	+	10.00	175.37	0.3933	10	! N
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y60397	-	9.00	140.13	2.82	128	! H
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y91334	+	9.00	132.07	2.48	409	! H
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y59353	+	9.00	131.68	2.47	433	! H
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y59354	+	9.00	131.68	2.47	433	! H
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y59355	+	9.00	131.20	2.45	464	! M
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y04995	+	9.00	129.65	2.39	580	! M
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W81666	+	9.00	129.65	2.39	580	! M
/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W32431	+	9.00	129.65	2.39	580	! M
/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W32363	+	9.00	129.65	2.39	580	! M
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W64303	+	9.00	129.39	2.38	602	! M
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y39105	+	8.00	116.95	24.50	288	! A
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y38968	+	8.00	116.74	24.42	297	! H
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80102	+	8.00	115.45	23.92	358	! H
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W98445	+	8.00	114.90	23.72	387	! H
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W89850	+	8.00	113.97	23.37	443	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W89977	+	8.00	113.84	23.32	451	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:R77434	+	8.00	113.09	23.04	503	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:Y55721	+	8.00	112.43	22.80	553	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y41277	+	8.00	112.43	22.80	553	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y43260	+	8.00	112.43	22.80	553	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R51355	+	8.00	112.39	22.79	556	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y53675	-	8.00	100.11	18.76	3262	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y53676	-	8.00	100.11	18.75	3264	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W39792	-	7.00	122.06	333.33	11	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA1991.DAT:R14324	-	7.00	117.91	312.10	20	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y20222	-	7.00	117.57	310.43	21	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R63153	+	7.00	115.83	301.95	27	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W46589	+	7.00	115.58	300.75	28	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y12582	-	7.00	110.77	278.65	56	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R58674	-	7.00	108.56	269.05	77	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y21088	+	7.00	107.95	266.49	84	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P90408	+	7.00	107.63	265.12	88	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1995.DAT:R80528	+	7.00	107.17	263.21	94	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y60272	+	7.00	107.03	262.60	96	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y60288	+	7.00	107.03	262.60	96	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1992.DAT:R29888	+	7.00	106.54	260.57	103	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y44878	+	7.00	106.54	260.57	103	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W98318	+	7.00	106.21	259.21	108	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1992.DAT:R26087	+	7.00	105.90	257.92	113	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1992.DAT:R26147	+	7.00	105.90	257.92	113	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1993.DAT:R38722	+	7.00	105.90	257.92	113	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W41979	+	7.00	105.90	257.92	113	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W76993	+	7.00	105.90	257.92	113	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W58853	+	7.00	105.77	257.43	115	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W38606	+	7.00	105.71	257.18	116	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W22886	+	7.00	105.42	255.99	121	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y50973	+	7.00	105.36	255.76	122	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y50975	+	7.00	105.36	255.76	122	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75703	+	7.00	105.03	254.41	128	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W64530	+	7.00	104.17	250.94	145	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y81598	+	7.00	103.61	248.75	157	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56766	+	7.00	102.94	246.11	173	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y36820	+	7.00	102.44	244.15	186	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y48501	+	7.00	102.18	243.16	193	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75704	+	7.00	101.87	241.95	202	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75705	+	7.00	101.87	241.95	202	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75706	+	7.00	101.87	241.95	202	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75707	+	7.00	101.87	241.95	202	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75205	+	7.00	101.83	241.81	203	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y12916	+	7.00	101.73	241.42	206	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y12969	+	7.00	101.50	240.54	213	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R47240	+	7.00	101.50	240.54	213	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80963	+	7.00	101.30	239.80	219	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P93648	+	7.00	101.30	239.80	219	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:R06438	+	7.00	101.27	239.68	220	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W69515	+	7.00	101.21	239.44	222	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75486	+	7.00	101.21	239.44	222	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W01819	+	7.00	100.70	237.51	239	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y63762	+	7.00	100.19	235.62	257	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y23988	+	7.00	100.19	235.31	260	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y23987	+	7.00	100.03	235.02	263	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W62627	+	7.00	99.90	234.53	268	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:W88281	+	7.00	99.72	233.87	275	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80960	+	7.00	99.43	232.77	287	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y56056	+	7.00	99.43	232.77	287	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y57853	+	7.00	99.43	232.77	287	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y57865	+	7.00	99.36	232.50	290	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80961	+	7.00	99.33	232.41	291	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1990.DAT:R03340	-	7.00	99.17	231.81	298	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W86809	-	7.00	99.01	231.22	305	-
/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:W62268	+	7.00	99.01	231.22	305	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y57057	+	7.00	98.80	230.48	314	+
/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y91560	+	7.00	98.80	230.48	314	+

						Path
H	314	230.48	98.80	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y91672
S	316	230.32	98.76	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W88749
H	320	230.00	98.67	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80962
E	323	229.76	98.61	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56767
E	330	229.22	98.46	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1992.DAT:R29926
A	330	229.22	98.46	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W58564
S	337	228.69	98.31	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1992.DAT:R26874
S	337	228.69	98.31	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1993.DAT:R35044
H	337	228.69	98.31	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:W77184
H	341	228.39	98.23	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y80999
S	343	228.25	98.19	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W27804
H	343	228.25	98.19	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W48667
H	344	228.17	98.17	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W48668
M	361	226.96	97.84	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80959
S	362	226.89	97.82	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P93567
S	362	226.89	97.82	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P90419
H	366	226.62	97.74	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y40499
M	373	226.15	97.61	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y04940
S	377	225.88	97.54	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:W56774
S	379	225.75	97.50	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P90508
S	380	225.69	97.48	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P90509
S	380	225.69	97.48	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P90510
S	380	225.69	97.48	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P90511
H	380	225.69	97.48	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P90512
E	382	225.56	97.44	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80958
M	385	225.36	97.39	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y73428
S	391	224.98	97.28	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:R97659
M	391	224.98	97.28	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y50335
H	394	224.79	97.23	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W51408
H	394	224.79	97.23	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1995.DAT:Y15193
M	395	224.73	97.21	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:R66920
S	395	224.73	97.21	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:W01952
H	395	224.73	97.21	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W88282
E	397	224.60	97.18	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:R66923
H	397	224.60	97.18	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:W01955
H	398	224.54	97.16	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1995.DAT:R66921
H	398	224.54	97.16	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:W01953
M	399	224.48	97.14	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1995.DAT:R66922
E	399	224.48	97.14	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:W01954
S	417	223.39	96.84	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1992.DAT:R27809
E	417	223.39	96.84	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1993.DAT:R37683
E	417	223.39	96.84	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1993.DAT:R37895
H	417	223.39	96.84	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1993.DAT:R41345
M	423	223.04	96.74	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:R77401
E	425	222.92	96.70	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:R75800
H	431	222.58	96.61	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56778
H	431	222.58	96.61	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56776
S	436	222.30	96.53	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y04941
H	450	221.52	96.31	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W62781
H	450	221.52	96.31	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W98654
H	451	221.47	96.29	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y40498
E	458	221.09	96.19	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56769
H	458	221.09	96.19	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56772
H	458	221.09	96.19	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56773
S	458	221.09	96.19	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W56777
H	466	220.67	96.07	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y06442
A	476	220.16	95.92	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y53635
C	480	219.95	95.86	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W63698
H	480	219.95	95.86	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W22849
H	480	219.95	95.86	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y80956
S	480	219.95	95.86	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1995.DAT:Y15191
g	491	219.41	95.70	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y70769
S	491	219.41	95.70	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y70770
g	507	218.63	95.48	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1993.DAT:R37682
H	509	218.54	95.45	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y81602
H	525	217.80	95.24	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R58659
H	529	217.61	95.19	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W73473

					Path
532	217.48	95.15	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y06443
534	217.39	95.12	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W25769
546	216.86	94.97	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W39788
547	216.81	94.95	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W40112
552	216.60	94.89	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R58661
553	216.55	94.88	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R58660
553	216.55	94.88	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R58662
565	216.04	94.73	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W96316
571	215.79	94.66	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W96315
598	214.70	94.33	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W82547
601	214.58	94.30	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:W03562
601	214.58	94.30	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W44291
604	214.46	94.27	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W42400
615	214.03	94.14	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W62779
615	214.03	94.14	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:W62780
640	213.10	93.86	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:Y57301
653	212.63	93.72	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:Y35021
698	211.07	93.26	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:Y91524
703	210.91	93.21	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W23599
705	210.84	93.19	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W29513
712	210.61	93.12	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1997.DAT:W30749
739	209.75	92.87	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y56971
752	209.35	92.74	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:R97199
842	206.76	91.96	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1989.DAT:P93712
854	206.44	91.86	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W37991
854	206.44	91.86	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y39872
916	204.85	91.38	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R45779
939	204.29	91.20	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y67882
961	203.77	91.04	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:R49039
971	203.54	90.97	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y44717
1036	202.09	90.52	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W53280
1074	201.29	90.27	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W53279
1095	200.86	90.14	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:W80359
1182	199.18	89.61	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y33496
1374	195.91	88.56	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W72225
1374	195.91	88.56	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W69753
1384	195.75	88.51	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W72224
1396	195.56	88.45	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W72117
1396	195.56	88.45	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:W72039
1445	194.82	88.21	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R58809
1694	191.44	87.11	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W40109
1719	191.13	87.01	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1996.DAT:R92100
1832	189.80	86.57	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y58575
1967	188.32	86.07	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1993.DAT:R33547
1968	188.31	86.07	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W63139
2138	186.60	85.49	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1995.DAT:R72607
2163	186.36	85.41	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:R71003
2243	185.62	85.09	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y14589
2250	185.55	85.14	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W67738
2261	185.45	85.11	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1995.DAT:Y14587
2268	185.39	85.09	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y14588
2476	183.61	84.48	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W71007
2510	183.33	84.38	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W72204
3119	179.00	82.87	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:Y08898
3139	178.87	82.83	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1999.DAT:Y39299
3170	178.68	82.76	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y50976
3256	178.15	82.58	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y77179
3782	175.24	81.54	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:W97246
4472	171.06	80.37	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R47861
4544	170.74	80.26	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1994.DAT:R60517
4544	170.74	80.26	7.00	+	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:R58659
5069	168.60	79.51	7.00	-	/cgn2_2/gcgdata/geneseq/geneseqp/AA1998.DAT:W52846

seq_name: /cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y83150

```
seq_documentation_block:
ID   Y83150 standard; Protein; 465 AA.
XX
AC   Y83150;
XX
DT   24-JUL-2000 (first entry)
XX
DE   NGSP polypeptide of Neisseria gonorrhoeae.
```

```
XX
KW   NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW   cellular matrix; adherence domain; ligand; detection; diagnosis;
KW   screening; probe; primer; prophylaxis; therapy.
XX
OS   Neisseria gonorrhoeae.
XX
PN   WO20001213-A1.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US20070.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ, Harris AM;
XX
DR   WPI: 2000-237782/20.
DR   N-PSDB: Z93414.
XX
PT   Non-cytosolic NGSP polypeptide and polynucleotide sequence from
PT   Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT   infections
XX
PS   Claim 5; Page 61-62; 68pp; English.
XX
CC   The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC   Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC   domains for extracellular matrix proteins. Using the NGSP polypeptide
CC   as a vaccine produces antibodies which inhibit binding of N.
CC   gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC   subsequent invasion. The NGSP polypeptide and its peptide fragments
CC   can be used to immunise an animal and produce an immune response.
CC   They can also be used as ligands to detect antibodies elicited in
CC   response to Neisseria infections and also as antigens or immunogens
CC   for inducing Neisseria-specific antibodies which are useful in
CC   immunoassays to detect Neisseria in biological specimens. Nucleotides
CC   encoding NGSP or its fragments can be used as probes to identify
CC   Neisseria in biological specimens by hybridization or polymerase
CC   chain reaction amplification. The NGSP polypeptide can also be used
CC   in screening assays to identify agents and compounds which useful as
CC   diagnostic, prophylactic or therapeutic agents against Neisseria
CC   infection.
XX
SQ   Sequence   465 AA;

alignment_scores:
     Quality:  465.00                Length:      465
     Ratio:      1.000               Gaps:          0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y83150

Align seg 1/1 to: Y83150   from: 1   to: 465
```

```
 51 SerMetLeuLeuProAspPheAlaGlnLeuValGlnSerGluGlyProAl  67

201 AGTCGTCAATATTCAGGAGCAGCCCCGCGCCCCGCGCACCCAAAACGGCAGCG  250
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 67 aValValAsnIleGlnAlaAlaProAlaProArgThrGlnAsnGlySerG  84

251 GCAATGCCGAAACCGATTCCGACCCGTTGCCGACCAGGAGGACCCGTTCTAC  300
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 84 lyAsnAlaGluThrAspSerAspProLeuAlaAspSerAspPProPheTyr 100

301 GAATTTTTCAAACGCCTCGTCCCGAACATGCCCGAAATCCCCAAGAAGA  350
    |||||||||||||||||||||||||||||||||||||||||||||||||
101 GluPhePheLysArgLeuValProAsnMetProGluIleProGlnGluGl 117

351 AGCAGATGACGGCGGATGAACTTCGGTTCGGGCTTCATCATCAGCAAAA  400
    |||||||||||||||||||||||||||||||||||||||||||||||||
117 uAlaAspAspGlyGlyLeuAsnPheGlySerGlyPheIleIleSerLysA 134

401 ACGGCTACATCTGACCAATACCCACGTCGTTGCCGGATGGGCAGTATC  450
    ||||||||||||||||||||||||||||||||||||||||||||||||
134 snGlyTyrIleLeuThrAsnThrHisValValAlaGlyMetGlySerIle 150

451 AAAGTCCTGCTCAACGACCAAGCGCGAATATACCCGCAAACTCATCGGTTC  500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 LysValLeuLeuAsnAspLysArgGluTyrThrArgLysLeuIleGlySe 167

501 GGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCAACGGAAGAGC  550
    |||||||||||||||||||||||||||||||||||||||||||||||||
167 rAspValGlnSerAspValAlaLeuLeuLysIleAspAlaThrGluGluL 184

551 TACCCGTCGTCAAAATCGGCAATCCCAAAAATTGAAACCGGGCGAATGG  600
    |||||||||||||||||||||||||||||||||||||||||||||||||
184 euProValValLysIleGlyAsnProLysAsnLeuLysProGlyGluTrp 200

601 GTCGCTGCCATCGGCGCCCCTTCGGCTTGACAACAGCGTGACCGCCGG  650
    |||||||||||||||||||||||||||||||||||||||||||||||||
201 ValAlaAlaIleGlyAlaProPheGlyPheAspAsnSerValThrAlaGl 217

651 CATCGTGTCCGCAAGGCAGAAGCCTGCCAACGAAGCTACACACCCT  700
    ||||||||||||||||||||||||||||||||||||||||||||||
217 yIleValSerAlaLysGlyArgSerLeuProAsnGluSerTyrThrProP 234

701 TCATCCAAACCGACGTTGCCATCAATCCGGCAATTCCGGCGGCCGCTG  750
    ||||||||||||||||||||||||||||||||||||||||||||||||
234 heIleGlnThrAspValAlaIleAsnProGlyAsnSerGlyGlyProLeu 250

751 TTCAACTTAAAAGGACAGGTCGTCGGCCATCAATTCGCAAATATACAGCCG  800
    |||||||||||||||||||||||||||||||||||||||||||||||||||
251 PheAsnLeuLysGlyGlnValValGlyIleGlnSerGlnIleTyrSerAr 267

801 CAGCGGCCGGATTCATGGGCATCTCCTTTGCCATCCCGATTGACGTTGCCA  850
    ||||||||.||||||||||||||||||||||||||||||||||||||||||
267 gSerGlyGlyPheMetGlyIleSerPheAlaIleProIleAspValAlaM 284

851 TGAAATGTCGCCCGAACAGCTGAAAAACACCGGCAAAGTCAACGGGACAA  900
    |||||||||||||||||||||||||||||||||||||||||||||||||
284 etAsnValAlaGluProAsnThrGlyLysValGlnArgGlyGln 300

901 CTGGGCGTGATTATTCAGGAAGTATCCTACGGTTTGGCACAGTCGTTCGG  950
    |||||||||||||||||||||||||||||||||||||||||||||||||
301 LeuGlyValIleIleGlnGluValSerTyrGlyLeuAlaGlnSerPheGl 317

951 TCTTGGATAAAGCCAGCGGCGCAATTGATTCCCAAAATCCTTCCCGCAGCC 1000
    |||||||||||||||||||||||||||||||||||||||||||||||||||
317 yLeuAspLysAlaSerGlyAlaLeuIleAlaLysIleLeuProGlySerP 334

1001 CCGCAGAACGTGCCGGCCTGCAGGCGGCGACATCGTCTCAGCCTCGAC 1050
     ||||||||||||||||||||||||||||||||||||||||||||||||
334 roAlaGluArgValGluGlnIleAlaGlyAspIleValLeuSerLeuAsp 350

1051 GGCGGAGAAATACGTTCTTCCGGCGACCTTCCGTCATGGTCGGCGCCAT 1100
     ||||||||||||||||||||||||||||||||||||||||||||||||
351 GlyGlyGluIleArgSerGlyAspLeuProValMetValGlyAlaAlaI 367
```

```
  1 GTGTTCAAAAAATACCAATACTTCGGCTTTGGCGGCCACTGTGTGCCGCCTT  50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ValPheLysLysTyrGlnTyrPheAlaLeuAlaLeuAlaLeuCysAlaAlaLe 17

 51 GCTGGCAGGCTGCGAAAAGGCAGGCAGCTTTTCGGTGCCGGACAAAAAAG 100
    |||||||||||||||||||||||||||||||||||||||||||||||||
 17 uLeuAlaGlyCysGluLysAlaGlyAlaGlySerPhePheGlyAlaAspLysLysG 34

101 AAGCATCCTTCGTAGAACGCATCGAACACACCAAAGACGACGGCAGTGTC 150
    |||||||||||||||||||||||||||||||||||||||||||||||||
 34 luAlaSerPheValGluArgIleGluHisThrLysAspAspGlySerVal 50

151 AGTATGCTGCTGCCCGACTTTGCCCAACTGGTTCAAAGCGAAGGCCCCGGC 200
    |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
1101 TACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGCGCAAAGGCGAAGAAA 1150
     |||||||||||||||||||||||||||||||||||||||||||||||||
 367 eThrProGlyLysGluValSerLeuGlyValTrpArgLysGlyGluGluI 384

1151 TCACAATCAAAGCCAAGCTGGGCAACGCGCCGAGCATACCGGCGCATCA 1200
     |||||||||||||||||||||||||||||||||||||||||||||||||
 384 leThrIleLysAlaLysLeuGlyAsnAlaAlaGluHisThrGlyAlaSer 400

1201 TCCAAAACAGATGAAGCCCCTACACCGAACAGCAATCCGGTACGTTCTC 1250
     |||||||||||||||||||||||||||||||||||||||||||||||||
 401 SerLysThrAspGluAlaProTyrThrGluGlnSerGlyThrPheSe 417

1251 GGTCGAATCCGCAGGCATTACCCTTCAGACACATACCGACAGCAGCGGCA 1300
     |||||||||||||||||||||||||||||||||||||||||||||||||
 417 rValGluSerAlaGlyIleThrLeuGlnThrHisThrAspSerSerGlyL 434

1301 AACACCTCGTCGTCGTACGGGTTCCGACGCGGCAGAACGCGCAGGCTTA 1350
     |||||||||||||||||||||||||||||||||||||||||||||||||
 434 ysHisLeuValValAlaArgValSerAspAlaAlaGluArgAlaGlyLeu 450

1351 AGGCACGGCGACGAAATCCTAGCCGTCAGGGCAAGTCCCCGTCAA 1395
     |||||||||||||||||||||||||||||||||||||||||||||
 451 ArgHisGlyAspGluIleLeuAlaValArgAlaSerProArgGln 465

seq_name: /cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75748

seq_documentation_block:
ID   Y75748 standard; Protein; 499 AA.
XX
AC   Y75748;
XX
DT   21-MAR-2000 (first entry)
XX
DE   Neisseria gonorrheae ORF 986 protein sequence SEQ ID NO:2968.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy.
XX
OS   Neisseria gonorrheae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   N-PSDB; Z54510.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 2; Page 1390; 1453pp; English.
XX
CC   253015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
```

```
CC   novel Neisseria meningitis and N. gonorrheae polynucleotides and
CC   polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence  499 AA.

alignment_scores:
             Quality:  451.00          Length:     451
               Ratio:    1.000           Gaps:        0
  Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y75748

Align seg 1/1  to: Y75748     from: 1  to: 499

  1 GTGTTCAAAAAATACCAATACTTCGCTTTGGCGGCCACTGTGTGCCGCCTT  50
    |||||||||||||||||||||||||||||||||||||||||||||||||
  1 ValPheLysLysTyrGlnTyrPheAlaLeuAlaAlaLeuCysAlaAlaLe  17

 51 GCTGGCAGGCTGCGGAAAAGGCAGGCAGCTTTTCGGTGCGACAAAAAAG  100
    |||||||||||||||||||||||||||||||||||||||||||||||||
 17 uLeuAlaGlyCysGluLysGlyAlaGlySerPhePheGlyAlaAspLysLysG  34

101 AAGCATCCTTCGTAGAACCGCATCGAACACACCAAAGACACGGCAGTGTC  150
    |||||||||||||||||||||||||||||||||||||||||||||||||
 34 luAlaSerPheValGluArgIleGluHisThrLysAspAspGlySerVal  50

151 AGTATGCTGCTGCCCGACTTGCCCAACTCGTTCAAAGCGAAGGCCCGGC  200
    |||||||||||||||||||||||||||||||||||||||||||||||||
 51 SerMetLeuLeuProAspPheAlaGlnLeuValGlnSerGluGlyProAl  67

201 AGTCGTCAATATTCAGGAGCCCCGCCCCGGCACCCAAAACGGCAGCGCG  250
    |||||||||||||||||||||||||||||||||||||||||||||||||
 67 aValValAsnIleGlnAlaAlaProAlaProArgThrGlnAsnGlySerG  84

251 GCAATGCCGAAACCGATTCCGACCCGTTCCGACAGCGACACCCGTTCTAC  300
    |||||||||||||||||||||||||||||||||||||||||||||||||
 84 lyAsnAlaGluThrAspSerAspProLeuAlaAspSerAspProPheTyr  100

301 GAATTTTCAAACGCCTCGTTCCCGGACATCCCCGAAATCCCCAAGAAGA  350
    |||||||||||||||||||||||||||||||||||||||||||||||||
101 GluPhePheLysArgLeuValProAsnMetProGlnIleProGlnGluGluI  117

351 AGCAGATGACGGCGGGATTGAACTTCGGTTCGGGCGTTCATCATCAGCAAAA  400
    |||||||||||||||||||||||||||||||||||||||||||||||||
117 uAlaAspAspGlyGlyLeuAsnPheGlySerGlyPheIleIleSerLysA  134

401 ACGGCTACATCCTGCACCAATACCCACGTCGTTGCCGGTATGGCCAGTATC  450
    |||||||||||||||||||||||||||||||||||||||||||||||||
134 snGlyTyrIleLeuThrAsnThrHisValValAlaGlyMetGlySerIle  150

451 AAAGTCCTGCTCAACGACACAGGCGCGAATATACCGCCAAACTCATCCGTTC  500
    |||||||||||||||||||||||||||||||||||||||||||||||||
151 LysValLeuLeuAsnAspLysArgGluTyrTyrThrAlaLysLeuGlySe  167

501 GGATGTCCAATCCGATGTCGCCCTTCGAAAATCGACGCAACGGAAGAGC  550
    |||||||||||||||||||||||||||||||||||||||||||||||||
167 rAspValGlnSerAspValAlaLeuLeuLysIleAspAlaThrGluGluL  184

551 TACCCGTCGTCAAAATGGCAATCCCAAAATTTGAAACCCGGCGAATGG  600
    |||||||||||||||||||||||||||||||||||||||||||||||||
184 euProValLysIleGlyAsnProLysAsnLeuLysProGlyGluTrp  200
```

```
XX
DE   NGSP polypeptide of Neisseria gonorrhoeae.
XX
KW   NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW   cellular matrix; adherence domain; ligand; detection; diagnosis;
KW   screening; probe; primer; prophylaxis; therapy.
XX
OS   Neisseria gonorrhoeae.
XX
PN   WO20012133-A1.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;       99WO-US20070.
XX
PR   01-SEP-1998;       98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-237782/20.
DR   N-PSDB; 293415.
XX
PT   Non-cytosolic NGSP polypeptide and polynucleotide sequence from
PT   Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT   infections
XX
PS   Claim 5; Page 63-64; 68pp; English.
XX
CC   The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC   Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC   domains for extracellular matrix proteins. Using the NGSP polypeptide
CC   as a vaccine produces antibodies which inhibit binding of N.
CC   gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC   subsequent invasion. The NGSP polypeptide and its peptide fragments
CC   can be used to immunise an animal and produce an immune response.
CC   They can also be used as ligands to detect antibodies elicited in
CC   response to Neisseria infections and also as antigens or immunogens
CC   for inducing Neisseria-specific antibodies which are useful in
CC   immunoassays to detect Neisseria in biological specimens. Nucleotides
CC   encoding NGSP or its fragments can be used as probes to identify
CC   Neisseria in biological specimens by hybridization or polymerase
CC   chain reaction amplification. The NGSP polypeptide can also be used
CC   in screening assays to identify agents and compounds which useful as
CC   diagnostic, prophylactic or therapeutic agents against Neisseria
CC   infection.
XX
SQ   Sequence   414 AA;
```

```
alignment_scores:
            Quality:  414.00                 Length:       414
            Ratio:      1.000                 Gaps:           0
Percent Similarity:  100.000       Percent Identity:   100.000
..
alignment_block:
US-09-388-090-3 x Y83151

Align seg 1/1   to: Y83151    from: 1   to: 414

154 ATGCTGCTGCCCGACTTTGCCCAACTGGTTCAAAGCGAAGGCCCGGCAGT 203
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 MetLeuLeuProAspPheAlaGlnLeuValGlnSerGluGlyProAlaVa 17

204 CGTCAATATTCAGGCAGCCCCGCCGCCCCGCCACCCAAAACGCAGCGGCA 253
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 17 lValAsnIleGlnAlaAlaProAlaProArgThrGlnAsnGlySerGlyA 34

254 ATGCCGAAAACCGATTCCGACCCCGTTGCCGACAGCGACCCGTTCTACGAA 303
    |||||||||||||||||||||||||||||||||||||||||||||||||
 34 snAlaGluThrAspSerAspProValAlaAspSerAspProPheTyrGlu 50
```

```
601 GTCGCTGCCATCGGCGGCGCCCTTCGGCTTGACAACAGCGTGACCGCCGG 650
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 ValAlaAlaIleGlyGlyAlaProGluGlyPheAspAsnSerValThrAlaGl 217

651 CATCGTGCTCCCCAAAGGCAGAGCCTGCCCACGAAAGCTACACACCCT 700
    ||||||||||||||||||||||||||||||||||||||||||||||||
217 yIleValSerAlaLysGlyArgSerLeuProAsnGluSerTyrThrProP 234

701 TCATCCAAACCGACGTTGCCATCAATCCGGCCAATTCCGGCCGCCCCGCTG 750
    ||||||||||||||||||||||||||||||||||||||||||||||||
234 heIleGlnThrAspValAlaIleAsnProGlyAsnSerGlyGlyProLeu 250

751 TTCAACTTAAAAGGACCAGGTCGTCGGCCATCAATTCGCAAATATACAGCCG 800
    ||||||||||||||||||||||||||||||||||||||||||||||||
251 PheAsnLeuLysGlyGlnValValGlyIleAsnSerGlnIleTyrSerAr 267

801 CAGCGGCGGATTCATGGGCCATCTCCTTTGCCATCCCGATTGACGTTGCCA 850
    ||||||||||||||||||||||||||||||||||||||||||||||||
267 gSerGlyGlyPheMetGlyIleSerPheAlaIleProIleAspValAlaM 284

851 TGAATGTCGGCCGAACAGCTGAAAAAACACCGGCAAAGTCCACGCGGACAA 900
    ||||||||||||||||||||||||||||||||||||||||||||||||
284 etAsnValAlaGluGlnGlnLeuLysAsnThrGlyLysValGlnArgGlyIln 300

901 CTGGGCCTGATTATTCAGGAAGTATCCTACGGTTTGGCCACAGTCGTTCGG 950
    ||||||||||||||||||||||||||||||||||||||||||||||||
301 LeuGlyValIleIleGlnGluValSerTyrGlyLeuAlaGlnSerPheGl 317

951 TCTGGATAAAGCCAGGCGCCGGCCATTGATTGCCAAATCCTTCCCGGCAGCC 1000
    ||||||||||||||||||||||||||||||||||||||||||||||||
317 yLeuAspLysProLysAlaSerGlyAlaAlaLeuIleAlaLysIleLeuProGlySerP 334

1001 CCGCAGAACGTGCCGGCCTCCAGGCGGCGGACATCGTCCTCAGCCTCGAC 1050
     ||||||||||||||||||||||||||||||||||||||||||||||||
334 roAlaGluArgAlaGlyLeuGlnAlaLysGlnAlaAspIleValLeuSerLeuAsp 350

1051 GGCGGAGAAATACGTTCTTCCGGCGACCTTCCGTCATGGTCGGCGGCCAT 1100
     ||||||||||||||||||||||||||||||||||||||||||||||||
351 GlyGlyGluIleArgSerSerGlyAspLeuProValMetValGlyAlaIl 367

1101 TACCGCCGGGAAAAGAAGTCAGCCTCCGGCGTATGGCCGCAAAGGCGAAGAAA 1150
     ||||||||||||||||||||||||||||||||||||||||||||||||
367 eThrProGlyLysGluValSerLeuGlyValTrpArgLysGlyGluGlui 384

1151 TCACAATCAAAGCCAAGCTGGGCAACGCGCGCGAGCATACCGGCGATCA 1200
     ||||||||||||||||||||||||||||||||||||||||||||||||
384 leThrIleLysAlaLysLeuGlyAsnAlaAlaGluHisThrGlyAlaSer 400

1201 TCCAAAACAGATGAAGCGCCCCTACACCGCCAACAGCAATCCGGTACGTTCTC 1250
     ||||||||||||||||||||||||||||||||||||||||||||||||
401 SerLysThrAspGluAlaProTyrThrGluGlnGlnSerGlyThrPheSe 417

1251 GGTCGAATCCCCAGGCATTACCCTTCAGACACCATACCGACAGCGGCA 1300
     ||||||||||||||||||||||||||||||||||||||||||||||||
417 rValGluSerAlaGlyIleThrLeuGlnThrHisThrAspSerSerGlyL 434

1301 AACACCTCGTCGTCGTACGGGTTTCCGACGGGCAGAACGCGCAGGCTTA 1350
     ||||||||||||||||||||||||||||||||||||||||||||||||
434 ysHisLeuValValArgValSerAspAlaAlaGluArgAlaGlyGlyLeu 450

1351 AGG 1353
     |||
451 Arg 451
```

```
seq_name: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:Y83151

seq_documentation_block:
ID   Y83151 standard; Protein: 414 AA.
XX
AC   Y83151;
XX
DT   24-JUL-2000  (first entry)
```

304 TTTTTCAAACGGCTCGTCCGAACATGCCGAAATCCCCAAGAAGAAGC 353
|||||||||||||||||||||||||||||||||||||||||||
51 PhePheLysArgLeuValProAsnMetProGluIleProGlnGluAl 67

354 AGATGACGGCGGATTGAACTTCGGTTCGGCTTCATCAGCAAAAACG 403
|||||||||||||||||||||||||||||||||||||||||||
67 aAspAspGlyGlyLeuAsnPheGlySerGlyPheIleIleSerLysAsnG 84

404 GCTACATCCTGACCAATACCCACGTCGTTGCCGTATGGCAGTATCAAA 453
|||||||||||||||||||||||||||||||||||||||||||
84 lyTyrIleLeuThrAsnThrHisValAlaAlaGlyMetGlySerIleLys 100

454 GTCCTGCTCAACGACAAGCGCGAATATACCGGCAAACTCATCGGTTCCGGA 503
|||||||||||||||||||||||||||||||||||||||||||
101 ValLeuLeuAsnAspLysArgGluTyrThrAlaLysLeuIleGlySerAs 117

504 TGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCAACGGAAGAGCTAC 553
|||||||||||||||||||||||||||||||||||||||||||
117 pValGlnSerAspValAlaLeuLeuLysIleAspAlaThrGluGluLeuP 134

554 CCGTCGTCAAAATCGGCAATCCCAAAAATTGAAACCGGGCGAATGGGTC 603
|||||||||||||||||||||||||||||||||||||||||||
134 roValValLysIleGlyAsnProLysAsnLeuLysProGlyGluTrpVal 150

604 GCTGCCATTCGGCGGCGCCCTTCGGCTTTGACAACAGCGTGACCCCGGCAT 653
|||||||||||||||||||||||||||||||||||||||||||
151 AlaAlaIleGlyGlyAlaProPheGlyPheAspAsnSerValThrAlaGlyIl 167

654 CGTGTCCGCCAAAGGCAGAAGCTGCCCAACGAAAGCTACACACCCTTCA 703
|||||||||||||||||||||||||||||||||||||||||||
167 eValSerAlaLysGlyArgSerLeuProAsnGluSerTyrThrProPheI 184

704 TCCAAACCGACGGTTGCCATCAATTCCGGGCGGCCCCGCTGTTC 753
|||||||||||||||||||||||||||||||||||||||||||
184 leGlnThrAspGlyCysHisGlnPheArgGlyProLeuPhe 200

754 AACTTAAAAGGACAGGTCGTCGGCATCAATTCCAAAATATACAGCCCGCAG 803
|||||||||||||||||||||||||||||||||||||||||||
201 AsnLeuLysGlyGlnValValGlyIleAsnSerGlnIleTyrSerArgSe 217

804 CGGCGGATTCATGGGCATCTCCTTTGCCATCCGATTGACGTTGCCATGA 853
|||||||||||||||||||||||||||||||||||||||||||
217 rGlyGlyPheMetGlyIleSerPheAlaIleProIleAspValAlaMetA 234

854 ATGTCGGCGAACAGTCGAAAACACCGGCAAAGTCCAACGCGGACAACTG 903
|||||||||||||||||||||||||||||||||||||||||||
234 snValAlaGluGluGlnSerLysAsnThrGlyLysValGlnArgGlyGlnLeu 250

904 GGCGTGATTATTCAGGAAGTATCCTACGGTTTGGCACAGTCGTTCGGGTCT 953
|||||||||||||||||||||||||||||||||||||||||||
251 GlyValIleIleGlnGluValSerTyrGlyLeuAlaGlnSerPheGlyLe 267

954 GGATAAAGCCAGCGGCGGCCATTGATTGCCAAAATCCTTCCCGGCAGCCCCG 1003
|||||||||||||||||||||||||||||||||||||||||||
267 uAspLysAlaSerGlyAlaLeuIleAlaLysIleLeuProGlySerProA 284

1004 CAGAACGTGCCGCCGCTGCAGGCGGCGGCGACATCGTCCTCAGCCTCGACGGC 1053
|||||||||||||||||||||||||||||||||||||||||||
284 laGluArgAlaAlaGlyLeuGlnAlaGlyAlaAspIleValLeuSerLeuAspGly 300

1054 GGAGAAATACGTTCTTCCGGCGACCTTCCCGTCATGGTCGGCGCCATTAC 1103
|||||||||||||||||||||||||||||||||||||||||||
301 GlyGluIleArgSerSerGlyAspLeuProValMetValGlyAlaIleTh 317

1104 GCCGGGAAAGAAGAGTCAGCCTCCGGTATGGCGCAAAGGCGAAGAAATCA 1153
|||||||||||||||||||||||||||||||||||||||||||
317 rProGlyLysGluValSerLeuGlyValTrpArgLysGlyGluGluIleT 334

1154 CAATCAAAGCCAAGTGGCAACGCCCGCGAGCATACCGGCGCATCATCC 1203
|||||||||||||||||||||||||||||||||||||||||||
334 hrIleLysAlaLysLeuGlyAsnAlaAlaGluHisThrGlyAlaSerSer 350

1204 AAAACAGTAGAGGCCCCTACACCGAACAGCAATCCGTACGTTCTCCGGT 1253
|||||||||||||||||||||||||||||||||||||||||||

351 LysThrAspGluAlaProTyrThrGlnGlnSerGlyThrPheSerVa 367

1254 CGAATCCGCAGGCATTACCCTTCAGACACATACCGACAGCAGGGCAAAC 1303
|||||||||||||||||||||||||||||||||||||||||||
367 lGluSerAlaGlyIleThrLeuGlnThrHisThrAspSerSerGlyLysH 384

1304 ACCTCGTCGTCGTACGGGTTTCCGACGCGGCAGAACGCGGCAGGCTTAAGG 1353
|||||||||||||||||||||||||||||||||||||||||||
384 isLeuValValArgValSerArgAspAlaSerAlaAlaGluArgAlaGlyLeuArg 400

1354 CACGGCGACGAAAATCCTAGCCGTCAGGGCAAGTCCCGTCAA 1395
|||||||||||||||||||||||||||||||||||||||||||
401 HisGlyAspGluAsnProSerArgGlnGlyLysSerArg 414

seq_name: /cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y52996

seq_documentation_block:
ID Y52996 standard; Protein; 370 AA.
XX
AC Y52996;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis BASB013-C protein sequence.
XX
KW Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW invasive bacterial disease; antibacterial.
XX
OS Neisseria meningitidis.
XX
PN WO9955872-A1.
XX
PD 04-NOV-1999.
XX
PF 20-APR-1999;   99WO-EP02765.
XX
PR 23-APR-1998;   98GB-0008734.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-052809/04.
DR N-PSDB; Z33308.
XX
PT Novel polynucleotides and polypeptides from Neisseria meningitis used
PT to prepare vaccines against bacterial infections -
XX
PS Example 2; Page 82-83; 94pp; English.
XX
CC The present sequence is a conserved BASB013-C polypeptide isolated
CC from Neisseria meningitidis. BASB013 polynucleotides and polypeptides
CC may be employed as research reagents and material for the discovery of
CC treatments and diagnostics for diseases, particularly human diseases.
CC They can be used for diagnosis of disease, staging of disease, or
CC determining response of an infectious organism to drugs. The
CC polynucleotides may be used as a source for hybridisation probes, and
CC for screening of genetic mutations, serotype, organism or strain
CC identification, identification of mutation in BASB013 sequences, and as
CC components of arrays which are useful for diagnostic and prognostic
CC purposes. The polypeptides can be used to produce antibodies. The
CC polypeptides can also be used in vaccine formulations, and to identify
CC agonists and antagonists. The polypeptides, antibodies, agonists and
CC antagonists (which are bacteristatic) are used for the treatment and
CC prevention of diseases such as upper respiratory tract infection,
CC invasive bacterial diseases such as bacteraemia and meningitis, and for
CC the development and screening of antibacterial drugs. They are also used
CC in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC on in-dwelling devices, or to extracellular proteins on wounds, and to
CC thus prevent tissue damage and/or block the normal progression of
CC pathogenesis in infections initiated other than by the implantation of

```
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence   370 AA;

alignment_scores:
        Quality:   127.00            Length:    127
         Ratio:    1.000              Gaps:      0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y52996      .

Align seq 1/1  to: Y52996  from: 1  to: 370

   583 TTGAAACCGGGCGAATGGGTCGCTGCCATCGGCGCGCCCTTCGGCTTTGA 632
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   195 LeuLysProGlyGluTrpValAlaAlaIleGlyAlaProPheGlyPheAs 211

   633 CAACAGCGTGACCGCCGGCATCGTGCCGCCAAAGGCAGAAGCCTGCCCA 682
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   211 pAsnSerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProA 228

   683 ACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCCATCAATCCGGGC 732
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   228 snGluSerTyrThrProPheIleGlnThrAspValAlaIleAsnProGly 244

   733 AATTCCGGCGCCCCGTGTTCAACTAAAAGGACAGTCGTCGGCATCAA 782
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   245 AsnSerGlyGlyProLeuPheAsnLeuLysGlyGlnValValGlyIleAs 261

   783 TTCGCAAATATACAGCCGCAGCGGCGGATTCATGGCATCTCCTTTGCCA 832
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   261 nSerGlnIleTyrSerArgSerGlyGlyPheMetGlyIleSerPheAlaI 278

   833 TCCCGATTGACGTTGCCATGAATGTCCCGAACAGCTGAAAAACACCGGC 882
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   278 leProIleAspValAlaMetAsnValAlaGluGlnLeuLysAsnThrGly 294

   883 AAAGTCCAACGCGGACAACTGGGCGTGATTATTCAGGAAGTATCCTACGG 932
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   295 LysValGlnArgGlyGlnLeuGlyValIleIleGlnGluValSerTyrGl 311

   933 TTTGGCACAGTCGTTCGGTCTCGGATAAAGCC 963
       ||||||||||||||||||||||||||||||||
   311 yLeuAlaGlnSerPheGlyLeuAspLysAla 321
```

seq_name: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:Y75749

```
seq_documentation_block:
ID   Y75749 standard; Protein; 499 AA.
XX
AC   Y75749;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 986 protein sequence SEQ ID NO:2970.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI   Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI   Tettelin H, Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   N-PSDB; Z54511.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 2; Page 1390; 1453pp; English.
XX
CC   253015 to 253536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC   novel Neisseria meningitis and N. gonorrheae polynucleotides and
CC   polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence   499 AA;

alignment_scores:
        Quality:   127.00            Length:    127
         Ratio:    1.000              Gaps:      0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y75749      .

Align seq 1/1  to: Y75749  from: 1  to: 499

   583 TTGAAACCGGGCGAATGGGTCGCTGCCATCGGCGCGCCCTTCGGCTTTGA 632
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   195 LeuLysProGlyGluTrpValAlaAlaIleGlyAlaProPheGlyPheAs 211

   633 CAACAGCGTGACCGCCGGCATCGTGCCGCCAAAGGCAGAAGCCTGCCCA 682
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   211 pAsnSerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProA 228

   683 ACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCCATCAATCCGGGC 732
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   228 snGluSerTyrThrProPheIleGlnThrAspValAlaIleAsnProGly 244

   733 AATTCCGGCGCCCCGTGTTCAACTAAAAGGACAGTCGTCGGCATCAA 782
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   245 AsnSerGlyGlyProLeuPheAsnLeuLysGlyGlnValValGlyIleAs 261

   783 TTCGCAAATATACAGCCGCAGCGGCGGATTCATGGCATCTCCTTTGCCA 832
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   261 nSerGlnIleTyrSerArgSerGlyGlyPheMetGlyIleSerPheAlaI 278

   833 TCCCGATTGACGTTGCCATGAATGTCCCGAACAGCTGAAAAACACCGGC 882
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   278 leProIleAspValAlaMetAsnValAlaGluGlnLeuLysAsnThrGly 294

   883 AAAGTCCAACGCGGACAACTGGGCGTGATTATTCAGGAAGTATCCTACGG 932
```

Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y52994

Align seg 1/1 to: Y52994 from: 1 to: 499

```
583  TTGAAACCGGGCGAATGGGTCGTGCCATCGGCGCGCCCTTCGGCTTGA  632
     ||||||||||||||||||||||||||||||||||||||||||||||||
195  LeuLysProGlyGluTrpValAlaAlaIleGlyAlaProPheGlyPheAs  211

633  CAACAGCTGACCGGCCGGCATCGTGTCCGCCAAAGCAGGAGCCTGCCCA  682
     ||||||||||||||||||||||||||||||||||||||||||||||||
211  pAsnSerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProA  228

683  ACGAAAGCTACACACCCTTCATCCAAACCGACGTTCCATCAATCCGGGC  732
     ||||||||||||||||||||||||||||||||||||||||||||||||
228  snGluSerTyrThrProPheIleGlnThrAspValAlaIleAsnProGly  244

733  AATTCCCGCGGCCCGCTGTTCAACTTAAAGGACAGGTCGTCGGCATCAA  782
     ||||||||||||||||||||||||||||||||||||||||||||||||
245  AsnSerGlyGlyProLeuPheAsnLeuLysGlyGlnValValGlyIleAs  261

783  TTCGCAAATATACAGCCCGCAGCCGGCGGATTCATGGGCATCTCCTTTGCCA  832
     |||||||||||||||||||||||||||||||||||||||||||||||
261  nSerGlnIleTyrSerArgSerGlyGlyPheMetGlyIleSerPheAlaI  278

833  TCCCGATTGACGTTGCCATGAATGTCGCCCAACAGCTGAAAAACACCGGC  882
     ||||||||||||||||||||||||||||||||||||||||||||||||
278  leProIleAspValAlaMetAsnValAlaGluGlnLeuLysAsnThrGly  294

883  AAAGTCCAACGCGGACAACTGGGCGTGATTATTCAGGAAGTATCCTACGG  932
     ||||||||||||||||||||||||||||||||||||||||||||||||
295  LysValGlnArgGlyGlnLeuGlyValIleIleGlnGluValSerTyrGl  311

933  TTTGGCACAGTCGTTCGGTCTGGATAAAGCC  963
     ||||||||||||||||||||||||||||||
311  yLeuAlaGlnSerPheGlyLeuAspLysAla  321
```

seq_name: /cgn2_2/gcgdata/geneseq/AA2000.DAT:Y52995

seq_documentation_block:
ID Y52995 standard; Protein; 499 AA.
XX
AC Y52995;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis strain H44/76 BASB013 protein sequence.
XX
KW Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW invasive bacterial disease; antibacterial.
XX
OS Neisseria meningitidis.
XX
PN WO995872-A1.
XX
PD 04-NOV-1999.
XX
PF 20-APR-1999; 99WO-EP02765.
XX
PR 23-APR-1998; 98GB-0008734.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-052809/04.
DR N-PSDB; Z33307.
XX
PT Novel polynucleotides and polypeptides from Neisseria meningitis used
PT to prepare vaccines against bacterial infections -

```
     ||||||||||||||||||||||||||||||||||||||||||||||||
295  LysValGlnArgGlyGlnLeuGlyValIleIleGlnGluValSerTyrGl  311

933  TTTGGCACAGTCGTTCGGTCTGGATAAAGCC  963
     ||||||||||||||||||||||||||||||
311  yLeuAlaGlnSerPheGlyLeuAspLysAla  321
```

seq_name: /cgn2_2/gcgdata/geneseq/AA2000.DAT:Y52994

seq_documentation_block:
ID Y52994 standard; Protein; 499 AA.
XX
AC Y52994;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis strain ATCC 13090 BASB013 protein sequence.
XX
KW Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW invasive bacterial disease; antibacterial.
XX
OS Neisseria meningitidis.
XX
PN WO995872-A1.
XX
PD 04-NOV-1999.
XX
PF 20-APR-1999; 99WO-EP02765.
XX
PR 23-APR-1998; 98GB-0008734.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-052809/04.
DR N-PSDB; Z33306.
XX
PT Novel polynucleotides and polypeptides from Neisseria meningitis used
PT to prepare vaccines against bacterial infections -
XX
PS Claim 3; Page 77-78; 94pp; English.
XX
CC The present sequence represents a BASB013 polypeptide isolated from
CC Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC employed as research reagents and material for the discovery of
CC treatments and diagnostics for diseases, particularly human diseases.
CC They can be used for diagnosis of disease, staging of disease, or
CC determining response of an infectious organism to drugs. The
CC polynucleotides may be used as a source for hybridisation probes, and
CC for screening of genetic mutations, serotype, organism or strain
CC identification, identification of mutation in BASB013 sequences, and as
CC components of arrays which are useful for diagnostic and prognostic
CC purposes. The polypeptides can be used to produce antibodies. The
CC polypeptides can also be used in vaccine formulations, and to identify
CC agonists and antagonists. The polypeptides, antibodies, agonists and
CC antagonists (which are bacteristatic) are used for the treatment and
CC prevention of diseases such as upper respiratory tract infection,
CC invasive bacterial diseases such as bacteraemia and meningitis, and for
CC the development and screening of antibacterial drugs. They are also used
CC in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC on in-dwelling devices, or to extracellular proteins on wounds, and to
CC thus prevent tissue damage and/or block the normal progression of
CC pathogenesis in infections initiated other than by the implantation of
CC in-dwelling devices or by other surgical techniques.
XX
SQ Sequence 499 AA;
```

alignment_scores:
Quality:  127.00    Length:  127
Ratio:  1.000    Gaps:  0

XX
PS   Claim 3; Page 79-81; 94pp; English.
XX
CC   The present sequence represents a BASB013 polypeptide isolated from
CC   Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC   employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence   499 AA;

alignment_scores:
         Quality: 127.00              Length:     127
           Ratio: 1.000                 Gaps:       0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y52995      ..

Align seg 1/1  to: Y52995   from: 1   to: 499

583 TTGAAACCGGGGAATGGGTCGCTGCCATCGGCGCGCCGCCCTTCGGCTTTGA 632
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
195 LeuLysProGlyGluTrpValAlaAlaIleGlyAlaProPheGlyPheAs 211

633 CAACAGCGTGACCCGCGGCATCGTTGTCCGCCAAAGGCAGAAGCCTGCCCA 682
    |||||||||||||||||||||||||||||||||||||||||||||||||||
211 pAsnSerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProA 228

683 ACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCCATCAATCCGGGC 732
    |||||||||||||||||||||||||||||||||||||||||||||||||
228 snGluSerTyrThrProPheIleGlnThrAspValAlaIleAsnProGly 244

733 AATTCCGGCGGCCCGCTGTTCAACTTAAAAGGACAGGTCGTCGGCATCAA 782
    |||||||||||||||||||||||||||||||||||||||||||||||||
245 AsnSerGlyGlyProLeuPheAsnLeuLysGlyGlnValValGlyIleAs 261

783 TTCGCAAATATACAGCCCGCAGCCGCGGATTCATGGCCATCCTTGCCA 832
    |||||||||||||||||||||||||||||||||||||||||||||||
261 nSerGlnIleTyrSerProGlnProArgIleMetGlyIleSerPheAlaI 278

833 TCCCGATTGACGTTGCCATGAATGTCGCCCAAACAGCTGAAAAACACCGGC 882
    |||||||||||||||||||||||||||||||||||||||||||||||||||
278 leProIleAspValAlaMetAsnValAlaGluGlnThrAlaGluLysHisArgGly 294

883 AAAGTCCAACGCGGACAACTGGGCGTGATTATTCAGGAAGTATCCTACGG 932
    ||||||||||||||||||||||||||||||||||||||||||||||||||
295 LysValGlnArgGlyGlnLeuGlyValIleIleGlnGluValSerTyrGl 311

933 TTTGGCACAGTCGTTCGGTCTCGGATAAAGCC 963
    ||||||||||||||||||||||||||||||||
311 yLeuAlaGlnSerPheGlyLeuAspLysAla 321

seq_name: /cgn2_2/gcqdata/geneseq/geneseqAA/AA2000.DAT:Y52993

---

seq_documentation_block:
ID   Y52993 standard; Protein; 499 AA.
XX
AC   Y52993;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Neisseria meningitidis strain ATCC 13090 BASB013 protein sequence.
XX
KW   Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW   antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW   invasive bacterial disease; antibacterial.
XX
OS   Neisseria meningitidis.
XX
PN   WO9955872-A1.
XX
PD   04-NOV-1999.
XX
PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-052809/04.
DR   N-PSDB; 233305.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections  -
XX
PS   Claim 5; Page 74-76; 94pp; English.
XX
CC   The present sequence represents a BASB013 polypeptide isolated from
CC   Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC   employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence   499 AA;

alignment_scores:
         Quality: 115.00              Length:     115
           Ratio: 1.000                 Gaps:       0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y52993      ..

Align seg 1/1  to: Y52993   from: 1   to: 499

619 CCCTTCGGCTTTGACAACAGCGTGACCCGCGGCATCGTGTCCGCCAAAGG 668
    ||||||||||||||||||||||||||||||||||||||||||||||||||

```
207 ProPheGlyPheAspAsnSerValThrAlaGlyIleValSerAlaLysGl 223
669 CAGAAGCCTGCCCAACGAAGCTACACACCCTTCATCCAAACCGACGTTG 718
    |||||||||||||||||||||||||||||||||||||||||||||||||
223 yArgSerLeuProAsnGluSerTyrThrProPheIleGlnThrAspValA 240

719 CCATCAATTCCGGGCAATTCCGGCGGCCCCGTGTTCAACTTAAAAGGACAG 768
    ||||||||||||||||||||||||||||||||||||||||||||||||||
240 lAIleAsnProGlyAsnSerGlyPheAsnLeuLysGlyGln 256

769 GTCGTCGGCATCAATTCGCAAATATACAGCCGCAGCGGCGGATTCATGGG 818
    |||||||||||||||||||||||||||||||||||||||||||||||||
257 ValValGlyIleAsnSerGlnIleTyrSerArgSerGlyGlyPheMetGl 273

819 CATCTCCTTGCCATCCGATTGACGTTGCCATGAATGTCGCCGAACAGC 868
    ||||||||||||||||||||||||||||||||||||||||||||||||
273 yIleSerPheAlaIleProIleAspValAlaMetAsnValAlaGluGlnL 290

869 TGAAAAACACCGGCAAAGTCCAACCGGACAACTGGGCGTGATTATTCAG 918
    |||||||||||||||||||||||||||||||||||||||||||||||||
290 euLysAsnThrGlyLysValGlnArgGlyGlnLeuGlyValIleIleGln 306

919 GAAGTATCCTACGGTTGGCACAGTCGTTCGGTCTGGATAAAGCC 963
    ||||||||||||||||||||||||||||||||||||||||||||
307 GluValSerTyrGlyLeuAlaGlnSerPheGlyLeuAspLysAla 321
```

seq_name: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:Y70409

seq_documentation_block:
ID Y70409 standard; Protein; 448 AA.
XX
AC Y70409;
XX
DT 03-JUL-2000 (first entry)
XX
DE Neisseria meningitidis NMASP protein-1.
XX
KW NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW Neisserial infection; meningitidis; septicaemia.
XX
OS Neisseria meningitidis.
XX
FH Key   Location/Qualifiers
FT Misc-difference 167
FT   /note= "Encoded by NTC"
XX
PN WO200012535-A2.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WO-US19663.
PR 01-SEP-1998; 98US-0098685.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ, Harris AM;
XX
DR WPI; 2000-256581/22.
DR N-PSDB; Z51533.
XX
PT Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT antibodies, useful in vaccines against infection -
XX
PS Claim 5; Page-; 75pp; English.
XX
CC The present sequence is the Neisseria meningitidis NMASP protein.
CC NMASP is a non-cytosolic protein, with antibacterial and
CC antiinflammatory activity. It shows sequence similarity to E. coli
CC DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC may be used for diagnosis, therapy or prophylaxis of Neisserial
CC infections such as, bacterial meningitidis and septicaemia.
CC Note: The protein represented in SEQ ID NO:2 of the specification is
CC erroneous. The present sequence is the decoded version of the nucleotide
CC represented in Z51533.
XX
SQ Sequence 448 AA;

alignment_scores:
        Quality: 103.00   Length: 103
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y70409

Align seg 1/1 to: Y70409 from: 1 to: 448

```
655 GTGTCCGCCAAAGCCAGAAGCCTGCCCAACGAAGCTACACACCCTTCAT 704
    |||||||||||||||||||||||||||||||||||||||||||||||||
168 ValSerAlaLysGlyArgSerLeuProAsnGluSerTyrThrProPheIl 184

705 CCAAACCGACGTTGCCATCAATCCGGCCAATTCCGGCGGCCCCGTGTTCA 754
    ||||||||||||||||||||||||||||||||||||||||||||||||||
184 eGlnThrAspValAlaIleAsnProGlyAsnSerGlyGlyProLeuPheA 201

755 ACTTAAAAGGACAGGTCGTCGGCATCAATTCGCAAATATACAGCGCGCAGC 804
    |||||||||||||||||||||||||||||||||||||||||||||||||
201 snLeuLysGlyGlnValValGlyIleAsnSerGlnIleTyrSerArgSer 217

805 GGCGGATTCATGGGCATCTCCTTTGCCATCCGATTGACGTTGCCATGAA 854
    |||||||||||||||||||||||||||||||||||||||||||||||||
218 GlyGlyPheMetGlyIleSerPheAlaIleProIleAspValAlaMetAs 234

855 TGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACCGGACAACTGG 904
    ||||||||||||||||||||||||||||||||||||||||||||||||
234 nValAlaGluGlnLeuLysAsnThrGlyLysValGlnArgGlyGlnLeuG 251

905 GCGTGATTATTCAGGAAGTATCCTACGGTTGGCACAGTCGTTCGGTCTG 954
    |||||||||||||||||||||||||||||||||||||||||||||||||
251 lyValIleIleGlnGluValSerTyrGlyLeuAlaGlnSerPheGlyLeu 267

955 GATAAAGCC 963
    |||||||||
268 AspLysAla 270
```

seq_name: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:Y70414

seq_documentation_block:
ID Y70414 standard; Protein; 475 AA.
XX
AC Y70414;
XX
DT 03-JUL-2000 (first entry)
XX
DE Neisseria meningitidis NMASP protein-3.
XX
KW NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW Neisserial infection; meningitidis; septicaemia.
XX
OS Neisseria meningitidis.
XX
FH Key   Location/Qualifiers
FT Misc-difference 19
FT   /note= "Encoded by GAA"
FT Misc-difference 35
FT   /note= "Encoded by GCC"
FT Misc-difference 61
FT   /note= "Encoded by GGC"
FT Misc-difference 111
FT   /note= "Encoded by AAC"

```
FT   Misc-difference 122
FT          /note= "Encoded by GCC"
FT   Misc-difference 171
FT          /note= "Encoded by AAT"
FT   Misc-difference 194
FT          /note= "Encoded by GGCATC"
FT   Misc-difference 298
FT          /note= "Encoded by AGC"
FT   Misc-difference 317
FT          /note= "Encoded by AGG"
FT   Misc-difference 364
FT          /note= "Encoded by GCC"
FT   Misc-difference 373
FT          /note= "Encoded by ACC"
FT   Misc-difference 410
FT          /note= "Encoded by AAA"
FT   Misc-difference 428
FT          /note= "Encoded by CAC"
FT   Misc-difference 436
FT          /note= "Encoded by AGG"
FT   Misc-difference 437
FT          /note= "Encoded by GCA"
FT   Misc-difference 438
FT          /note= "Encoded by AGT"
FT   Misc-difference 440..475
FT          /note= "Encoded by CGTCAA"
XX
PN   WO200012535-A2.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US19663.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-256581/22.
DR   N-PSDB; 251539.
XX
PT   Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT   antibodies, useful in vaccines against infection  -
XX
PS   Claim 5; Page 71-73; 75pp; English.
XX
CC   The present sequence is the Neisseria meningitidis NMASP protein.
CC   NMASP is a non-cytosolic protein, with antibacterial and
CC   antiinflammatory activity. It shows sequence similarity to E. coli
CC   DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC   antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC   anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC   may be used for diagnosis, therapy or prophylaxis of Neisserial
CC   infections such as, bacterial meningitidis and septicaemia.
XX
SQ   Sequence  475 AA;

alignment_scores:
       Quality: 103.00          Length:        103
       Ratio: 1.000             Gaps:            0
Percent Similarity: 100.000     Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y70414   ..

Align seq 1/1   to: Y70414   from: 1  to: 475

655 GTGTCCGCCAAAGGCAGAAGGCTGCCAACGAAAGCTACACACCTTCAT 704
    ||||||||||||||||||||||||||||||||||||||||||||||||
195 ValSerAlaLysGlyArgSerLeuProAsnGluSerTyrThrProPheI 211
```

```
705 CCAAACCGACGTTGCCATCAATCCGGCGCAATTCCGGCGCGGCCCGCTGTTCA 754
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
211 eGlnThrAspValAlaIleAsnProGlyAsnSerGlyGlyProLeuPheA 228

755 ACTTAAAAGGACAGGTCGTCGGCATCAATTCGCAAATATACAGCCGCAGC 804
    ||||||||||||||||||||||||||||||||||||||||||||||||||
228 snLeuLysGlyGlnValValGlyIleAsnSerGlnIleTyrSerArgSer 244

805 GGCGGATTCATGGCATCTCCTTTGCCATCCCGATTGACGTTGCCATGAA 854
    ||||||||||||||||||||||||||||||||||||||||||||||||||
245 GlyGlyPheMetGlyIleSerPheAlaIleProIleAspValAlaMetAs 261

855 TGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAACTGG 904
    ||||||||||||||||||||||||||||||||||||||||||||||||||
261 nValAlaGluGlnLeuLysAsnThrGlyLysValGlnAsnArgGlyGlnLeuG 278

905 GCGTGATTATTCAGGAAGTATCCTACGGTTGGCACAGTCGTTCGGTCTG 954
    ||||||||||||||||||||||||||||||||||||||||||||||||||
278 lyValIleIleGlnGluValSerTyrGlyLeuAlaGlnSerPheGlyLeu 294

955 GATAAAGCC 963
    |||||||||
295 AspLysAla 297

seq_name: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:Y70413

seq_documentation_block:
ID   Y70413 standard; Protein; 498 AA.
XX
AC   Y70413;
XX
DT   03-JUL-2000   (first entry)
XX
DE   Neisseria meningitidis NMASP protein-2.
XX
KW   NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW   anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW   Neisserial infection; meningitidis; septicaemia.
XX
OS   Neisseria meningitidis.
XX
FH   Key              Location/Qualifiers
FT   Misc-difference 8
FT          /note= "Encoded by TTC"
FT   Misc-difference 17
FT          /note= "Encoded by TTG"
FT   Misc-difference 22
FT          /note= "Encoded by GAA"
FT   Misc-difference 42
FT          /note= "Encoded by GAA"
FT   Misc-difference 58
FT          /note= "Encoded by GCC"
FT   Misc-difference 84
FT          /note= "Encoded by GGC"
FT   Misc-difference 134
FT          /note= "Encoded by AAC"
FT   Misc-difference 145
FT          /note= "Encoded by GCC"
FT   Misc-difference 194
FT          /note= "Encoded by AAT"
FT   Misc-difference 218
FT          /note= "Encoded by GGCATC"
FT   Misc-difference 321
FT          /note= "Encoded by AGC"
FT   Misc-difference 340
FT          /note= "Encoded by AGG"
FT   Misc-difference 387
FT          /note= "Encoded by GCC"
FT   Misc-difference 396
FT          /note= "Encoded by ACC"
FT   Misc-difference 433
FT          /note= "Encoded by AAA"
FT   Misc-difference 451
```

```
FT
FT   Misc-difference   459
FT                     /note= "Encoded by CAC"
FT   Misc-difference   460
FT                     /note= "Encoded by AGG"
FT   Misc-difference   461
FT                     /note= "Encoded by GCA"
FT   Misc-difference   463
FT                     /note= "Encoded by AGT"
FT   Misc-difference   464
FT                     /note= "Encoded by CGT"
FT   Misc-difference   465..498
FT                     /note= "Encoded by CAA"
FT                     /note= "Encoded by AGGGCAAGTCCCGTCAA"
XX
PN   WO200012535-A2.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US19663.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-256581/22.
DR   N-PSDB; 251538.
XX
PT   Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT   antibodies, useful in vaccines against infection  -
XX
PS   Claim 5; Page 70-71; 75pp; English.
XX
CC   The present sequence is the Neisseria meningitidis NMASP protein.
CC   NMASP is a non-cytosolic protein, with antibacterial and
CC   antiinflammatory activity. It shows sequence similarity to E. coli
CC   DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC   antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC   anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC   may be used for diagnosis, therapy or prophylaxis of Neisserial
CC   infections such as, bacterial meningitidis and septicaemia.
XX
SQ   Sequence    498 AA;

alignment_scores:
     Quality:    103.00          Length:      103
     Ratio:      1.000           Gaps:          0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y70413

Align seq 1/1   to: Y70413  from:  1  to: 498

655  GTGTCCGCCAAAGGCAGAAGCCTGCCAACGAAAGCTACACACCCTTCAT 704
     |||||||||||||||||||||||||||||||||||||||||||||||||
218  ValSerAlaLysGlyArgSerLeuProAsnGluSerTyrThrProPheIl 234

705  CCAAACCGACGTTGCCATCAATCCGGCAATTCCGGCGGCCGGCTGTTCA 754
     |||||||||||||||||||||||||||||||||||||||||||||||||
234  eGlnThrAspValAlaIleAsnProGlyAsnSerGlyGlyProLeuPheA 251

755  ACTTAAAAGGACCAGAGGTCGTCGGCATCAATTCGCAAATATACAGCCCCAGC 804
     |||||||||||||||||||||||||||||||||||||||||||||||||
251  snLeuLysGlyGlnValGlyIleAsnSerGlnIleTyrSerArgSer 267

805  GCGGATTCATGGGCATCTCCTTTGCCATCCGATTGACGTTGCCATGAA 854
     |||||||||||||||||||||||||||||||||||||||||||||||||
268  GlyGlyPheMetGlyIleSerPheAlaIleProIleAspValAlaMetAs 284

855  TGTCGCCAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAACTGG 904
```

```
              ||||||||||||||||||||||||||||||||||||||||||||
         284  nValAlaGluGlnLeuLysAsnThrGlyLysValGlnArgGlyGlnLeuG 301

         905  GCGTGATTATTCAGGAAGTATCCTACGGTTGGCACAGTCGTTCGGTCTG 954
              ||||||||||||||||||||||||||||||||||||||||||||
         301  lyValIleIleGlnGluValSerTyrGlyLeuAlaGlnSerPheGlyLeu 317

         955  GATAAAGCC 963
              |||||||||
         318  AspLysAla 320

seq_name: /cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y75750

seq_documentation_block:
ID   Y75750 standard; Protein; 499 AA.
XX
AC   Y75750;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 986 protein sequence SEQ ID NO:2972.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   N-PSDB; 254512.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 2; Page 1392-1393; 1453pp; English.
XX
CC   Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC   novel Neisseria meningitis and N. gonorrheae polynucleotides and
CC   polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence    499 AA;
```

alignment_scores:
```
                Quality:  103.00              Length:         103
                  Ratio:    1.000              Gaps:             0
Percent Similarity: 100.000      Percent Identity: 100.000
```

alignment_block:
US-09-388-090-3 x Y75750   ..

Align seg 1/1 to: Y75750 from: 1 to: 499

```
 655 GTGTCCGCAAAGGCAGAGAGCCTGCCCAACGAAAGCTACACACCCTTCAT 704
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 219 ValSerAlaLysGlyArgSerLeuProAsnGluSerTyrThrProPheIl 235

 705 CCAAACCGACGTTGCCATCAATCCGGCAATTCCGGCGCCGCTGTTCA 754
     ||||||||||||||||||||||||||||||||||||||||||||||
 235 eGlnThrAspValAlaIleIleAsnProGlyAsnSerGlyGlyProLeuPheA 252

 755 ACTTAAAAGGACAGCTGCTCGGCATCAATTCGCAAATATACAGCCGCAGC 804
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 252 snLeuLysGlyGlnValValGlyIleAsnSerGlnIleTyrSerArgSer 268

 805 GGCGGATTCATGGGCATCTCCTTTGCCATCCCGATTGACGTTGCCATGAA 854
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 269 GlyGlyPheMetGlyIleSerPheAlaIleProIleAspValAlaMetAs 285

 855 TGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAACTGG 904
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 285 nValAlaGluGlnLeuLysAsnThrGlyLysValGlnArgGlyGlnLeuG 302

 905 GCGTGATTATTCAGGAGTATCCTACGGTTGGCACAGTCGTTCGGTCTG 954
     ||||||||||||||||||||||||||||||||||||||||||||||||
 302 lyValIleIleGlnGluValSerTyrGlyLeuAlaGlnSerPheGlyLeu 318

 955 GATAAAGCC 963
     |||||||||
 319 AspLysAla 321
```

seq_name: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:Y83152

seq_documentation_block:
ID Y83152 standard; Peptide; 51 AA.
XX
AC Y83152;
XX
DT 24-JUL-2000 (first entry)
XX
DE NGSP peptide of Neisseria gonorrhoeae.
XX
KW NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW cellular matrix; adherence domain; ligand; detection; diagnosis;
KW screening; probe; primer; prophylaxis; therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO20012133-A1.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WO-US20070.
XX
PR 01-SEP-1998; 98US-0098685.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ, Harris AM;
XX
DR WPI; 2000-237782/20.
DR N-PSDB; Z93416.
XX
PD 04-NOV-1999.
XX
PT Non-cytosolic NGSP polypeptide and polynucleotide sequence from

---

PT Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT infections
XX
PS Claim 5; Page 64-65; 68pp; English.
XX
CC The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC domains for extracellular matrix proteins. Using the NGSP polypeptide
CC as a vaccine produces antibodies which inhibit binding of N.
CC gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC subsequent invasion. The NGSP polypeptide and its peptide fragments
CC can be used to immunise an animal and produce an immune response.
CC They can also be used as ligands to detect antibodies elicited in
CC response to Neisseria infections and also as antigens or immunogens
CC for inducing Neisseria specific antibodies which are useful in
CC immunoassays to detect Neisseria in biological specimens. Nucleotides
CC encoding NGSP or its fragments can be used as probes to identify
CC Neisseria in biological specimens by hybridization or polymerase
CC chain reaction amplification. The NGSP polypeptide can also be used
CC in screening assays to identify agents and compounds which useful as
CC diagnostic, prophylactic or therapeutic agents against Neisseria
CC infection.
XX
SQ Sequence   51 AA;

alignment_scores:
```
                Quality:   51.00              Length:          51
                  Ratio:    1.000              Gaps:             0
Percent Similarity: 100.000      Percent Identity: 100.000
```

alignment_block:
US-09-388-090-3 x Y83152   ..

Align seg 1/1 to: Y83152 from: 1 to: 51

```
  1 GTGTTCAAAAAATACCAATACTTCGCTTTGGCGGCACTGTGTGCCGCCTT 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ValPheLysLysTyrGlnTyrPheAlaLeuAlaAlaLeuCysAlaAlaLe 17

 51 GCTGGCAGGCTGCGAAAAGGCAGGCAGCTTTTTCGGTGCGGACAAAAAAG 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 17 uLeuAlaGlyCysGluLysAlaGlySerPhePheGlyAlaAspLysLysG 34

101 AAGCATCCTTCGTAGAACGCATCGAACACCACCAAAGACGACGGCCAGTGTC 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 34 luAlaSerPheValGlnArgIleGluHisThrLysAspAspGlySerVal 50

151 AGT 153
    |||
 51 Ser 51
```

seq_name: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:Y52997

seq_documentation_block:
ID Y52997 standard; Protein; 129 AA.
XX
AC Y52997;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis variable BASB013-V protein sequence.
XX
KW Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW invasive bacterial disease; antibacterial.
XX
OS Neisseria meningitidis.
XX
PN WO9955872-A1.
XX
PD 04-NOV-1999.
XX

DE   Neisseria meningitidis NMASP protein N-terminal fragment.
XX
KW   NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW   anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW   Neisserial infection; meningitidis; septicaemia.
XX
OS   Neisseria meningitidis.
XX
PN   WO200012535-A2.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US19663.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ, Harris AM;
XX
DR   WPI: 2000-256581/22.
DR   N-PSDB: 251542.
XX
PT   Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT   antibodies, useful in vaccines against infection -
XX
PS   Disclosure; Page 74; 75pp; English.
XX
CC   The present sequence is the N-terminal fragment of Neisseria
CC   meningitidis NMASP protein. NMASP is a non-cytosolic protein, with
CC   antibacterial and antiinflammatory activity. It shows sequence similarity
CC   to E. coli DegP (HtrA) protein. NMASP proteins can be used as ligands to
CC   detect antibodies elicited in response to N. meningitidis infections.
CC   Cytotoxic anti-NMASP antibodies can be used as vaccines. NMASP proteins
CC   and DNA may be used for diagnosis, therapy or prophylaxis of Neisserial
CC   infections such as, bacterial meningitidis and septicaemia.
XX
SQ   Sequence   51 AA;

alignment_scores:
        Quality: 19.00   Length: 19
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y70415

Align seq 1/1 to: Y70415   from: 1   to: 51

67 AAGGCAGGCAGCTTTTCGGTGCGGACAAAAAGAAGCATCCTTCGTAGA 116
    ||||||||||||||||||||||||||||||||||||||||||||||||
23 LysAlaGlySerPhePheGlyAlaAspLysLysGluAlaSerPheValGl 39

117 ACGCATC 123
     |||||||
 39 uArgIle 41

PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
XX
PA   (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI: 2000-052809/04.
DR   N-PSDB: 233309.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections -
XX
PS   Example 2: Page 83-84; 94pp; English.
XX
CC   The present sequence is a variable BASB013-V polypeptide isolated
CC   from Neisseria meningitidis. BASB013 polynucleotides and polypeptides
CC   may be employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence   129 AA;

alignment_scores:
        Quality: 36.00   Length: 36
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Y52997

Align seq 1/1 to: Y52997   from: 1   to: 129

1192 GGCGCATCATCCAAAACAGATGAGCCCCTACACCGAACAGCAATCCGG 1241
      ||||||||||||||||||||||||||||||||||||||||||||||||
  28 GlyAlaSerSerLysThrAspGluAlaProTyrThrGluGlnSerGl 44

1242 TACGTTCTCGGTCGAATCCGCAGGCATTACCCTTCAGACACATACCGACA 1291
      ||||||||||||||||||||||||||||||||||||||||||||||||
  44 yThrPheSerValGluSerAlaGlyIleThrLeuGlnThrHisThrAspS 61

1292 GCAGCCGC 1299
      ||||||||
  61 erSerGly 63

seq_name: /cgn2_2/gcgdata/geneseq/geneseqp/AA2000.DAT:Y70415

seq_documentation_block:
ID   Y70415 standard; Protein; 51 AA.
XX
AC   Y70415;
XX
DT   03-JUL-2000 (first entry)
XX

us-09-388-090-3.oligo7.rag

```
QM of: US-09-388-090-3 to: Issued_Patents_AA:*     out_format : pfs

Date: Mar 22, 2001 12:05 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09388090/runat_20032001_092950_25795/app_query.fasta_1.1474
-DB=Issued_Patents_AA  -QFMT=fastan -SUFFIX=oligo7.rai
-GAPOP=4.500 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000
-XGAPEXT=60.000 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000
-YGAPEXT=60.000 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=oligo -TRANS=human40.cdi -LIST=1000 -DOCALIGN=200
-THR_SCORE=quality -THR_MIN=7 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MAXLEN=2000000000
-USER=US09388090_@CGN1_1_19 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-388-090-3
Query length: 1395
Database: Issued_Patents_AA:*
Database sequences: 17472
Database length: 17957048
Search time (sec): 23.850000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000
```

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	!	Documentation
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-142-469-3	+	8.00	133.43	16.49	27	!	
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-142-469-2	+	8.00	117.13	12.13	297	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-245-294-8	+	8.00	113.55	11.33	503	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-474-499-8	+	8.00	113.55	11.33	503	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-307-279A-8	+	8.00	113.55	11.33	503	!	
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT:US95-06211-8	+	8.00	113.55	11.33	503	!	
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-083-351-2	+	8.00	112.90	11.20	553	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-445-586-7	+	8.00	112.89	11.20	554	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-445-586-2	+	8.00	112.87	11.20	556	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-669-721-18	+	7.00	122.14	172.16	16	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-320-373-79	-	7.00	119.59	164.08	16	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-320-373-80	-	7.00	119.59	164.08	16	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-906-769-163	+	7.00	119.59	164.08	16	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-472-040A-7	-	7.00	118.43	160.50	19	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-906-616-163	+	7.00	116.04	153.43	27	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-639-075A-163	+	7.00	116.04	153.43	27	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-297-395-7	-	7.00	116.04	153.43	27	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-164-318-7	-	7.00	107.56	130.74	94	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-471-341-7	-	7.00	106.31	127.69	113	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-461-566-7	-	7.00	106.31	127.69	113	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-297-395-10	-	7.00	106.31	127.69	113	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-466-860-7	-	7.00	106.31	127.69	113	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-468-853-6	-	7.00	103.41	120.90	173	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-468-857-6	+	7.00	101.22	115.99	239	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-773-870-6	-	7.00	99.17	111.60	323	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-468-853-6	+	7.00	99.13	111.51	325	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-310-357-6	-	7.00	99.02	111.29	330	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-468-852-6	+	7.00	99.02	111.29	330	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-468-857-6	+	7.00	99.02	111.29	330	!	
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-108-020-53	+	7.00	98.12	109.41	377	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-923-454A-27	+	7.00	97.87	108.90	391	!	
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT:US95-15696-2	-	7.00	97.87	108.90	391	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-742-040A-8	+	7.00	97.82	108.79	394	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-097-938-2	+	7.00	97.80	108.75	395	!	

Documentation	Strd			Len
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-097-938-5	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-476-000-2	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-476-000-5	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-472-840-2	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-472-840-5	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-476-976-2	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-476-976-5	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-474-410-2	+	7.00	97.80	108.75	395
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-474-410-5	+	7.00	97.77	108.68	397
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-476-000-63	+	7.00	97.77	108.68	397
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-472-840-63	+	7.00	97.77	108.68	397
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-476-976-63	+	7.00	97.77	108.68	397
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-474-410-63	+	7.00	97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-097-938-4	+	7.00	97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-097-938-6	+	7.00	97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-476-000-4	+	7.00	97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-476-000-6	+	7.00	97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-472-840-4	+	7.00	97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-472-840-6	+	7.00	97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-476-976-4	+		97.75	108.65	398
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-476-976-6	+		97.75	108.65	398
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-474-410-4	+	7.00	97.73	108.61	399
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-474-410-6	+	7.00	97.73	108.61	399
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-476-000-61	+	7.00	97.73	108.61	399
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-472-840-61	+	7.00	97.73	108.61	399
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-476-976-61	+	7.00	97.43	108.00	417
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-474-410-61	-	7.00	97.34	107.80	423
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-682-847-6	+	7.00	97.21	107.54	431
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-923-454A-6	+	7.00	97.21	107.54	431
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-985-950-14	+	7.00	97.21	107.38	436
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-985-950-20	+	7.00	96.80	106.71	458
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-923-454A-29	+	7.00	96.80	106.71	458
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-923-454A-25	+	7.00	96.48	106.07	480
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-923-454A-31	+	7.00	95.87	104.86	525
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-923-454A-18	+	7.00	95.87	104.86	525
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-811-897A-23	+	7.00	95.60	104.63	534
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-855-213-23	+	7.00	95.59	104.33	546
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-691-814B-8	+	7.00	95.59	104.31	547
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-669-721-3	+	7.00	95.53	104.18	552
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-494-168-7	-	7.00	95.53	104.18	552
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-811-897A-27	+	7.00	95.51	104.16	553
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-855-213-27	+	7.00	95.51	104.16	553
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-811-897A-25	+	7.00	95.51	104.16	553
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-811-897A-25	+	7.00	95.51	104.16	553
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-855-213-25	+	7.00	95.51	104.16	553
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-855-213-29	+	7.00	95.37	103.87	565
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-892-770-6	-	7.00	95.30	103.73	571
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-892-770-5	+	7.00	94.95	103.05	601
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-606-288-10	+		94.95	103.05	601
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-606-288-7	+	7.00	94.95	103.05	601
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-968-752B-4	-	7.00	94.79	102.75	615
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-968-752B-2	+	7.00	94.79	102.75	615
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-347-483-7	+	7.00	93.97	101.17	694
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-347-483-10	-	7.00	92.56	98.51	854
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-985-950-16	+	7.00	92.36	98.15	879
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-985-950-18	+	7.00	91.25	96.10	1036
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-559-397-31	-	7.00	91.00	95.66	1074
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-928-692-17	-	7.00	91.00	95.66	1074
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-554-612C-1	+	7.00	89.00	92.11	1442
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-399-411-4	+	7.00	89.00	92.11	1442
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-015-986A-3	+	7.00	88.99	92.09	1445
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-446-363-3	+	7.00	88.99	92.09	1445
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-015-986A-2	+	7.00	87.91	90.23	1694
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-446-363-2	+	7.00	87.81	90.06	1719
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-494-168-2	-	7.00	87.81	90.06	1719
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-459-568-4	+	7.00	87.81	90.06	1719
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-399-411-4	-	7.00	87.37	89.32	1832
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-516-859A-4	+	7.00	86.89	88.51	1968
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-335-409-4	-	7.00	86.89	88.51	1968
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-07-745-206A-7	+	7.00	86.89	88.51	1968
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-455-543A-45	+				
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-223-305C-45	+				
```

```
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-311-363-7 +    7.00    86.89   88.51    1968 !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-276-967-2 -    7.00    85.33   85.94    2476 !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-149-097D-35 +          85.24   85.79    2509 !
/cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-036-987A-4 -    7.00    83.65   83.26    3170 !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-469-486-52 -   7.00    81.20   79.09    4544 !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-469-658-52 -   7.00    81.20   79.09    4544 !
```

seq_name: /cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-142-469-3

seq_documentation_block:
; Sequence 3, Application US/09142469
; Patent No. 6140094
; GENERAL INFORMATION:
; APPLICANT: LOFFLER, Fridolin
; APPLICANT: JUNGSCHAFFER, Gerald
; APPLICANT: KHANH, Quoc Nguyen
; APPLICANT: SCHUSTER, Erwin
; APPLICANT: SPROSSLER, Bruno
; APPLICANT: WOLF, Sabine
; TITLE OF INVENTION: PROTEIN WITH PHOSPHOLIPASE ACTIVITY
; FILE REFERENCE: 015200-056
; CURRENT APPLICATION NUMBER: US/09/142,469
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/EP98/00081
; EARLIER FILING DATE: 1998-01-08
; EARLIER APPLICATION NUMBER: DE 197 01 348.1
; EARLIER FILING DATE: 1997-01-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Aspergillus lysophospholipase
US-09-142-469-3

alignment_scores:
        Quality: 8.00   Length: 8
          Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-09-142-469-3

Align seg 1/1 to: US-09-142-469-3 from: 1 to: 27

```
        25 GCTTTGGCGGCACTGTGTGCCGCC 48
           |||||||||||||||||||||||||
        12 AlaLeuAlaAlaLeuCysAlaAla 19
```

seq_name: /cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-142-469-2

seq_documentation_block:
; Sequence 2, Application US/09142469
; Patent No. 6140094
; GENERAL INFORMATION:
; APPLICANT: LOFFLER, Fridolin
; APPLICANT: JUNGSCHAFFER, Gerald
; APPLICANT: KHANH, Quoc Nguyen
; APPLICANT: SCHUSTER, Erwin
; APPLICANT: SPROSSLER, Bruno
; APPLICANT: WOLF, Sabine
; TITLE OF INVENTION: PROTEIN WITH PHOSPHOLIPASE ACTIVITY
; FILE REFERENCE: 015200-056
; CURRENT APPLICATION NUMBER: US/09/142,469
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/EP98/00081
; EARLIER FILING DATE: 1998-01-08
; EARLIER APPLICATION NUMBER: DE 197 01 348.1
; EARLIER FILING DATE: 1997-01-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

; LENGTH: 297
; TYPE: PRT
; ORGANISM: Aspergillus lysophospholipase
US-09-142-469-2

alignment_scores:
        Quality: 8.00   Length: 8
          Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-09-142-469-2

Align seg 1/1 to: US-09-142-469-2 from: 1 to: 297

```
        25 GCTTTGGCGGCACTGTGTGCCGCC 48
           |||||||||||||||||||||||||
        12 AlaLeuAlaAlaLeuCysAlaAla 19
```

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-245-294-8

seq_documentation_block:
; Sequence 8, Application US/08245294
; Patent No. 5644047
; GENERAL INFORMATION:
; APPLICANT: Anderson, Burt E.
; APPLICANT: Regnery, Russell L.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DIAGNOSING
; TITLE OF INVENTION: ROCHALIMAEA HENSELAE AND ROCHALIMAEA QUINTANA
; TITLE OF INVENTION: INFECTION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG,P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,294
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.612
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-245-294-8

alignment_scores:
        Quality: 8.00   Length: 8
          Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-08-245-294-8

US 0955947

Align seg 1/1 to: US-08-245-294-8 from: 1 to: 503

640 GTGACCGCCGGCATCGTGTCCGCC 663
    ||||||||||||||||||||||||
215 ValThrAlaGlyIleValSerAla 222

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-474-499-8

seq_documentation_block:
; Sequence 8, Application US/08474499
; Patent No. 5693776
; GENERAL INFORMATION:
;   APPLICANT: Anderson, Burt E.
;   APPLICANT: Regnery, Russell L.
;   TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
;   TITLE OF INVENTION: DIAGNOSING ROCHALIMAEA HENSELAE
;   TITLE OF INVENTION: AND ROCHALIMAEA QUINTANA INFECTION
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: NEEDLE & ROSENBERG, P.C.
;   STREET: 127 Peachtree Street, Suite 1200
;   CITY: Atlanta
;   STATE: Georgia
;   COUNTRY: USA
;   ZIP: 30303
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/474,499
;   FILING DATE: 07-JUN-1995
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/245,294
;   FILING DATE: 18-MAY-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Spratt, Gwendolyn D.
;   REGISTRATION NUMBER: 36,016
;   REFERENCE/DOCKET NUMBER: 1414.612
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 404/688-0770
;   TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 503 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-474-499-8

alignment_scores:
        Quality: 8.00   Length: 8
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-08-474-499-8   ..

Align seg 1/1 to: US-08-474-499-8 from: 1 to: 503

640 GTGACCGCCGGCATCGTGTCCGCC 663
    ||||||||||||||||||||||||
215 ValThrAlaGlyIleValSerAla 222

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-307-279A-8

seq_documentation_block:
; Sequence 8, Application US/08307279A
; Patent No. 5736347
; GENERAL INFORMATION:

;   APPLICANT: Anderson, Burt E.
;   APPLICANT: Regnery, Russell L
;   TITLE OF INVENTION: Nucleic Acids of Rochalimaea Henselae
;   TITLE OF INVENTION: and Methods and Compositions for Diagnosing Rochalimaea
;   TITLE OF INVENTION: Henselae and Rochalimaea Quintana Infection
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: NEEDLE & ROSENBERG, P.C.
;   STREET: 127 Peachtree Street, N.E., Suite 1200
;   CITY: Atlanta
;   STATE: Georgia
;   COUNTRY: USA
;   ZIP: 30303
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/307,279A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Spratt, Gwendolyn D.
;   REGISTRATION NUMBER: 36,016
;   REFERENCE/DOCKET NUMBER: 1414.624
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (404) 688-0770
;   TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 503 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-307-279A-8

alignment_scores:
        Quality: 8.00   Length: 8
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-08-307-279A-8   ..

Align seg 1/1 to: US-08-307-279A-8 from: 1 to: 503

640 GTGACCGCCGGCATCGTGTCCGCC 663
    ||||||||||||||||||||||||
215 ValThrAlaGlyIleValSerAla 222

seq_name: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US95-06211-8

seq_documentation_block:
; Sequence 8, Application PC/rUS9506211
; GENERAL INFORMATION:
;   APPLICANT:
;   TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;   TITLE OF INVENTION: ROCHALIMAEA HENSELAE AND ROCHALIMAEA QUINTANA INFECTION
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: NEEDLE & ROSENBERG, P.C.
;   STREET: 127 Peachtree Street, Suite 1200
;   CITY: Atlanta
;   STATE: Georgia
;   COUNTRY: USA
;   ZIP: 30303
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25

```
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US95/06211
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/245,294
;     FILING DATE: 18 MAY 1994
;     CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Spratt, Gwendolyn D.
;     REGISTRATION NUMBER: 36,016
;     REFERENCE/DOCKET NUMBER: 1414.6121
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 404/688-0770
;     TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 503 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
PCT-US95-06211-8

alignment_scores:
      Quality:        8.00         Length:        8
      Ratio:          1.000        Gaps:          0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x PCT-US95-06211-8  ..

Align seg 1/1  to: PCT-US95-06211-8  from: 1  to: 503

640 GTGACCGCCGGCGATCGTGTCCGCC 663
    |||||||||||||||||||||||||
215 ValThrAlaGlyIleValSerAla 222

seq_name: /cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-083-351-2

seq_documentation_block:
; Sequence 2, Application US/09083351
; Patent No. 6087107
; GENERAL INFORMATION:
;   APPLICANT: Sheffield, Val C.
;   APPLICANT: Alward, Wallace L.M.
;   APPLICANT: Stone, Edwin M.
;   APPLICANT: Nishimura, Darryl
;   APPLICANT: Patil, Shiva
;   TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR
;   TITLE OF INVENTION: CONGENITAL HEART DISEASE BASED ON A NOVEL HUMAN
;   TITLE OF INVENTION: TRANSCRIPTION FACTOR
;   NUMBER OF SEQUENCES: 22
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: FOLEY, HOAG & ELIOT LLP
;     STREET: One Post Office Square
;     CITY: Boston
;     STATE: MA
;     COUNTRY: USA
;     ZIP: 02109-2170
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/083,351
;     FILING DATE: 22-MAY-1998
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Arnold, Beth E.
;     REGISTRATION NUMBER: 35,430
;     REFERENCE/DOCKET NUMBER: UIA-029.02
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 617-832-1000
```

```
;     TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 553 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-09-083-351-2

alignment_scores:
      Quality:        8.00         Length:        8
      Ratio:          1.000        Gaps:          0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-09-083-351-2  ..

Align seg 1/1  to: US-09-083-351-2  from: 1  to: 553

599 GGGTCGCTGCCATCGGCGGCGCCCT 622
    |||||||||||||||||||||||||
275 GlySerLeuProSerAlaArgPro 282

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-445-586-7

seq_documentation_block:
; Sequence 7, Application US/08445586
; Patent No. 5627050
; GENERAL INFORMATION:
;   APPLICANT: Takeshita, Sunao
;   APPLICANT: Ito, Toshimi
;   APPLICANT: Otawara-Hamamoto, Yoko
;   APPLICANT: Amann, Egon
;   TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
;   TITLE OF INVENTION: Process for its Production
;   NUMBER OF SEQUENCES: 11
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;     ADDRESSEE: Dunner
;     STREET: 1300 I Street, N.W.
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20005-3314
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/445,586
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/111,887
;     FILING DATE: 26-AUG-1993
;     APPLICATION NUMBER: JP 230030/92
;     FILING DATE: 28-AUG-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 324034/92
;     FILING DATE: 03-DEC-1992
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Forman, David S.
;     REGISTRATION NUMBER: 33,694
;     REFERENCE/DOCKET NUMBER: 02481.1322-00000
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-408-4000
;     TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 554 amino acids
;     TYPE: amino acid
```

```
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-445-586-7

alignment_scores:
        Quality:          8.00              Length:        8
        Ratio:            1.000             Gaps:          0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-08-445-586-7    ..

Align seg 1/1  to: US-08-445-586-7  from: 1  to: 554

    217 GCAGCCCCGCCCCGGCGCACCCAA 240
        |||||||||||||||||||||||
     18 AlaAlaProAlaProArgThrGln 25

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-445-586-2

seq_documentation_block:
; Sequence 2. Application US/08445586
; Patent No. 5627050
; GENERAL INFORMATION:
;   APPLICANT: Takeshita, Sunao
;   APPLICANT: Ito, Toshimi
;   APPLICANT: Otawara-Hamamoto, Yoko
;   APPLICANT: Amann, Egon
;   TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
;   TITLE OF INVENTION: Process for its Production
;   NUMBER OF SEQUENCES: 11
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;     ADDRESSEE: Dunner
;     STREET: 1300 I Street, N.W.
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 2005-3314
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/445,586
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/111,887
;     FILING DATE: 26-AUG-1993
;     APPLICATION NUMBER: JP 230030/92
;     FILING DATE: 28-AUG-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 324034/92
;     FILING DATE: 03-DEC-1992
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Forman, David S.
;     REGISTRATION NUMBER: 33,694
;     REFERENCE/DOCKET NUMBER: 02481.1322-00000
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-408-4000
;     TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 556 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-445-586-2
```

```
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-445-586-2

alignment_scores:
        Quality:          8.00              Length:        8
        Ratio:            1.000             Gaps:          0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-08-445-586-2    ..

Align seg 1/1  to: US-08-445-586-2  from: 1  to: 556

    217 GCAGCCCCGCCCCGGCGCACCCAA 240
        |||||||||||||||||||||||
     18 AlaAlaProAlaProArgThrGln 25

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-669-721-18

seq_documentation_block:
; Sequence 18. Application US/08669721
; Patent No. 5834236
; GENERAL INFORMATION:
;   APPLICANT: Lamb et al., Christopher J.
;   TITLE OF INVENTION: NOVEL TRANSCRIPTION ENHANCER ELEMENT AND
;   TITLE OF INVENTION: TRANSCRIPTION FACTOR AND METHODS OF USE THEREFOR
;   NUMBER OF SEQUENCES: 22
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Fish & Richardson P.C.
;     STREET: 4225 Executive Square, Suite 1400
;     CITY: La Jolla
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 92037
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/669,721
;     FILING DATE: 27-JUN-1996
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Ellison, Eldora L.
;     REGISTRATION NUMBER: 39,967
;     REFERENCE/DOCKET NUMBER: 07251/014001
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 619/678-5070
;     TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 11 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: not relevant
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-669-721-18

alignment_scores:
        Quality:          7.00              Length:        7
        Ratio:            1.000             Gaps:          0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-08-669-721-18    ..

Align seg 1/1  to: US-08-669-721-18  from: 1  to: 11

    180 GGTTCAAAGCGAAGGCCCGGC 200
        |||||||||||||||||||||
      1 GlySerLysArgArgProGly 7

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-320-373-79
```

seq_documentation_block:
; Sequence 79, Application US/08320373
; Patent No. 5559025
; GENERAL INFORMATION:
;   APPLICANT: Ahorn, Horst
;   APPLICANT: Maurer-Fogy, Ingrid
;   APPLICANT: Sommergruber, Wolfgang
;   APPLICANT: Zophel, Andreas
;   APPLICANT: Blaas, Dieter
;   APPLICANT: Kuchler, Ernst
;   APPLICANT: Liebig, Hans-Dieter
;   APPLICANT: Skern, Timothy
;   TITLE OF INVENTION: Expression of Mature Proteinase 2A, the
;   TITLE OF INVENTION: Partial Purification Thereof and Preparation of Substrates
;   TITLE OF INVENTION: Having an Inhibitory Effect
;   NUMBER OF SEQUENCES: 91
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox
;   STREET: 1225 Connecticut Avenue, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20036
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/320,373
;   FILING DATE: 11-OCT-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/971,619
;   FILING DATE: 06-NOV-1992
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 466-0800
;   TELEFAX: (202) 833-8716
;   TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 79:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 16 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
US-08-320-373-79

alignment_scores:
       Quality: 7.00   Length: 7
         Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3/rev x US-08-320-373-79   ..
Align seg 1/1 to: US-08-320-373-79 from: 1 to: 16

     209 TTGACGACTGCCGGGCCTTCG 189
         |||||||||||||||||||||
       5 LeuThrThrAlaGlyProSer 11

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-320-373-80

seq_documentation_block:
; Sequence 80, Application US/08320373
; Patent No. 5559025
; GENERAL INFORMATION:
;   APPLICANT: Ahorn, Horst
;   APPLICANT: Maurer-Fogy, Ingrid
;   APPLICANT: Sommergruber, Wolfgang
;   APPLICANT: Zophel, Andreas
;   APPLICANT: Blaas, Dieter
;   APPLICANT: Kuchler, Ernst
;   APPLICANT: Liebig, Hans-Dieter
;   APPLICANT: Skern, Timothy
;   TITLE OF INVENTION: Expression of Mature Proteinase 2A, the
;   TITLE OF INVENTION: Partial Purification Thereof and Preparation of Substrates
;   TITLE OF INVENTION: Having an Inhibitory Effect
;   NUMBER OF SEQUENCES: 91
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox
;   STREET: 1225 Connecticut Avenue, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20036
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/320,373
;   FILING DATE: 11-OCT-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/971,619
;   FILING DATE: 06-NOV-1992
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 466-0800
;   TELEFAX: (202) 833-8716
;   TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 80:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 16 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
US-08-320-373-80

alignment_scores:
       Quality: 7.00   Length: 7
         Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3/rev x US-08-320-373-80   ..
Align seg 1/1 to: US-08-320-373-80 from: 1 to: 16

     209 TTGACGACTGCCGGGCCTTCG 189
         |||||||||||||||||||||
       5 LeuThrThrAlaGlyProSer 11

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-637-759B-412

seq_documentation_block:
; Sequence 412, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
;   APPLICANT: David William Holden
;   TITLE OF INVENTION: Identification of Genes
;   NUMBER OF SEQUENCES: 501
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Patrea L. Pabst
;   STREET: 2800 One Atlantic Center
;   STREET: 1201 West Peachtree Street
;   CITY: Atlanta
;   STATE: Georgia
;   COUNTRY: USA
;   ZIP: 30309-3450
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30

```
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/637,759B
;   FILING DATE:  03-MAY-1996
;   CLASSIFICATION:  435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  PCT/GB95/02875
;   FILING DATE:  11-DEC-1995
;   CLASSIFICATION:  435
; ATTORNEY/AGENT INFORMATION:
;   NAME:  Pabst, Patrea L.
;   REGISTRATION NUMBER:  31,284
;   REFERENCE/DOCKET NUMBER:  RPMS 101
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (404) 873-8794
;   TELEFAX:  (404) 873-8795
; INFORMATION FOR SEQ ID NO: 412:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  16 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;   HYPOTHETICAL:  NO
US-08-637-759B-412

alignment_scores:
     Quality:       7.00          Length:       7
     Ratio:         1.000         Gaps:         0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-08-637-759B-412   from: 1   to: 16

Align seg 1/1   to: US-08-637-759B-412   from: 1   to: 16

    790 ATATACAGCCCAGCGGCCGGA 810
        |||||||||||||||||||||
      2 IleTyrSerArgSerGlyGly 8

seq_name: /cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-871-355A-412

seq_documentation_block:
; Sequence 412, Application US/08871355A
; Patent No. 6015569
; GENERAL INFORMATION:
;   APPLICANT:  David William Holden
;   TITLE OF INVENTION:  Identification of Genes
;   NUMBER OF SEQUENCES:  501
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Patrea L. Pabst
;     STREET:  2800 One Atlantic Center
;     STREET:  1201 West Peachtree Street
;     CITY:  Atlanta
;     STATE:  Georgia
;     COUNTRY:  USA
;     ZIP:  30309-3450
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/871,355A
;     FILING DATE:  09-JUN-1997
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  PCT/GB95/02875
;     FILING DATE:  11-DEC-1995
;     CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Pabst, Patrea L.
;     REGISTRATION NUMBER:  31,284
;     REFERENCE/DOCKET NUMBER:  RPMS 101 CON
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (404) 873-8794
;     TELEFAX:  (404) 873-8795
; INFORMATION FOR SEQ ID NO:  412:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  16 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;   HYPOTHETICAL:  NO
US-08-871-355A-412

alignment_scores:
     Quality:       7.00          Length:       7
     Ratio:         1.000         Gaps:         0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x US-08-871-355A-412   ..

Align seg 1/1   to: US-08-871-355A-412   from: 1   to: 16

    790 ATATACAGCCCAGCGGCCGGA 810
        |||||||||||||||||||||
      2 IleTyrSerArgSerGlyGly 8

seq_name: /cgn2_6/ptodata/2/iaa/6_COMB.pep:US-08-297-395-7

seq_documentation_block:
; Sequence 7, Application US/08297395A
; Patent No. 6039947
; GENERAL INFORMATION:
;   APPLICANT:  Howard L. Weiner
;   APPLICANT:  David A. Hafler
;   TITLE OF INVENTION:  PEPTIDES DERIVED FROM IMMUNODOMINANT
;   TITLE OF INVENTION:  EPITOPES OF MYELIN BASIC PROTEIN
;   FILE REFERENCE:  1010/05723US3
;   CURRENT APPLICATION NUMBER:  US/08/297,395A
;   CURRENT FILING DATE:  1994-08-11
;   EARLIER APPLICATION NUMBER:  08/059,189
;   EARLIER FILING DATE:  1993-05-06
;   EARLIER APPLICATION NUMBER:  07/502,559
;   EARLIER FILING DATE:  1990-03-30
;   EARLIER APPLICATION NUMBER:  PCT/US88/02139
;   EARLIER FILING DATE:  1988-06-24
;   EARLIER APPLICATION NUMBER:  07/065,734
;   EARLIER FILING DATE:  1987-06-24
;   NUMBER OF SEQ ID NOS:  84
;   SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH:  19
;   TYPE:  PRT
;   ORGANISM:  Homo sapiens
US-08-297-395-7

alignment_scores:
     Quality:       7.00          Length:       7
     Ratio:         1.000         Gaps:         0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3/rev x US-08-297-395-7   ..

Align seg 1/1   to: US-08-297-395-7   from: 1   to: 19

    534 GATTTCAGAAGGGCGACATC 514
        ||||||||||||||||||||
      8 AspPheGlnLysGlyAspIle 14
```

us-09-388-090-3.oligo7.rai

Fri Mar 23 07:26:25 2001

OM of: US-09-388-090-3 to: PIR_66:*   out_format : pfs

Date: Mar 22, 2001 12:05 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool/US09388090/runat_20032001_092950_25808/app_query.fasta_1.1474
-DB=PIR_66 -QFMT=fastan -SUFFIX=oligo7.rpr -GAPOP=4.500
-GAPEXT=0.100 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR_SCORE=quality
-THR_MIN=7 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09388090_@CGN1_1_71 -NCPU=6 -ICPU=3
-LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-388-090-3
Query length: 1395
Database: PIR_66:*
Database sequences: 195891
Database length: 67900655
Search time (sec): 40.680000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
pir2:B81914	+	127.00	2307.34	2.8e-121	499	! probable periplasmic serine pro
pir2:F83550	+	19.00	319.86	1.5e-10	474	! serine proteinase MucD precurso
pir2:F70048	+	11.00	177.56	0.0267	224	! serine proteinase Do homolog yv
pir2:S74643	+	11.00	173.83	0.0245	394	! proteinase hhoA (EC 3.4.-.-) -
pir2:B71284	+	11.00	173.76	0.0245	398	! probable periplasmic serine pro
pir2:S75445	+	11.00	173.47	0.0243	416	! proteinase hhoB (EC 3.4.-.-) -
pir2:A69643	+	11.00	173.09	0.0241	441	! probable periplasmic serine pro
pir2:S77538	+	11.00	172.97	0.0241	449	! serine proteinase Do, heat-shoc
pir1:B70426	+	11.00	172.92	0.0240	452	! serine proteinase (EC 3.4.21.-)
pir2:F72359	+	11.00	172.91	0.0240	453	! periplasmic serine proteinase D
pir2:G81528	+	11.00	172.82	0.0239	459	! periplasmic serine proteinase D
pir2:G72011	+	11.00	172.42	0.0237	488	! serine proteinase, HtrA/DegQ/De
pir2:B81728	+	11.00	172.42	0.0237	488	! do serine proteinase - Chlamydo
pir2:H71465	+	11.00	172.30	0.0236	497	! serine proteinase, HtrA/DegQ/De
pir2:T35460	-	9.00	138.93	2.88	295	! probable do serine proteinase -
pir2:T35287	+	9.00	137.58	2.79	362	! integral membrane protein - Sti
pir2:B83089	+	9.00	137.10	2.76	389	! probable secreted proteinase -
pir2:B70092	+	9.00	136.92	2.75	400	! AlgW protein PA4446 [imported]
pir2:T45448	+	9.00	136.11	2.69	452	! serine proteinase Do homolog YY
pir2:C70821	-	9.00	135.94	2.68	464	! probable serine proteinase (EC
pir2:H71465	-	9.00	135.80	2.68	474	! htrA-like protein - Brucella ab
pir1:A82581	+	9.00	135.26	2.64	514	! periplasmic proteinase XF2241 [
pir2:T02457	-	9.00	133.04	2.51	720	! hypothetical protein F4I18.12 [
pir2:T12203	+	9.00	132.42	2.47	790	! hypothetical protein L1231.2 [i
pir2:T02824	-	9.00	132.16	2.46	822	! transcription factor Vp1 - comm
pir2:E64040	-	8.00	121.70	31.32	247	! hypothetical protein HI1701 - H
pir2:T35304	+	8.00	120.82	30.69	282	! probable peptidase - Streptomyc
pir2:T22250	+	8.00	120.46	30.43	298	! hypothetical protein F45H11.1 -
pir2:E71023	+	8.00	120.41	30.40	300	! probable transcription initiati
pir2:S34116	+	8.00	120.41	30.40	300	! transcription factor IIB - Pyro
pir2:T46683	+	8.00	120.41	30.40	300	! transcription factor IIB [valid
pir2:E75110	-	8.00	120.41	30.40	300	! transcription initiation factor
pir2:D72693	-	8.00	120.18	30.24	311	! hypothetical protein APE0964 -
pir2:E71726	-	8.00	119.23	29.58	359	! hypothetical protein RP161 - Ri
pir1:G49964	-	8.00	119.10	29.49	366	! photosynthetic reaction center
pir2:T50891	+	8.00	119.10	29.49	366	! cytochrome subunit of photosynt
pir2:I48361	+	8.00	118.05	28.79	429	! gene GEG-154 protein - mouse
pir1:C64647	+	8.00	117.84	28.65	443	! serine proteinase (EC 3.4.21.-)

Documentation	Sequence	Strd	Orig	ZScore	EScore	Len
hypothetical protein HI0852	pir2:B64160	-	8.00	117.55	28.45	463
serine proteinase (proteinas	pir2:F81329	+	8.00	117.42	28.37	472
proteinase DO - Helicobacter	pir2:H71936	+	8.00	117.36	28.33	476
heat shock protein XF0285 [i	pir2:D82826	+	8.00	117.30	28.29	481
glutamate--tRNA ligase (EC 6	pir1:SYBSES	+	8.00	117.19	28.22	489
probable membrane protein -	pir2:T36336	+	8.00	116.08	27.51	578
transcription initiation fac	pir1:F64397	-	8.00	115.08	26.87	673
hypothetical protein YPR090W	pir2:S69074	+	8.00	114.49	26.51	736
pyrroloquinoline quinone syn	pir2:S5S241	+	8.00	113.70	26.03	829
hypothetical protein F17J16	pir2:T47784	+	8.00	111.76	24.89	1112
probable membrane protein Y	pir2:S69048	+	8.00	110.85	24.37	1276
agrin - electric ray (Disco	pir2:T43060	+	8.00	110.59	24.23	1328
copia-like polyprotein - Ar	pir2:T47492	+	8.00	110.41	24.13	1363
glutamate synthase (NADPH)	pir2:T34869	+	8.00	109.72	23.75	1514
L1 protein - human papillomav	pir2:E44906	+	7.00	116.20	447.22	35
hypothetical protein VCA0437	pir2:F82259	!	7.00	115.16	436.55	41
protamine phi-3.1 - Californi	pir2:S04941	!	7.00	114.54	430.39	45
protamine phi-3.2 - Californi	pir2:S10544	!	7.00	114.54	430.39	45
protamine phi-3.3 - Californi	pir2:S10545	!	7.00	114.54	430.39	45
hypothetical protein VCA0381	pir2:B82466	!	7.00	112.75	412.96	59
hypothetical protein VCA0412	pir2:C82461	!	7.00	112.75	412.96	59
hypothetical protein VCA0435	pir2:D82459	!	7.00	112.75	412.96	59
hypothetical protein VCA0407	pir2:B82465	!	7.00	112.64	411.90	60
hypothetical protein NMA0301	pir2:E82025	!	7.00	111.53	401.45	71
hypothetical protein T20K14_2	pir2:T51537	!	7.00	111.53	401.45	71
hypothetical protein NMB2130	pir2:C81003	!	7.00	110.74	394.21	80
T-cell receptor beta-chain V-	pir2:S52779	!	7.00	110.66	393.46	81
hypothetical protein - Deinoc	pir2:E75417	!	7.00	110.19	389.19	87
cuticle structural protein PC	pir2:S78005	!	7.00	109.82	385.89	92
hypothetical protein PA2143 [	pir2:H83378	!	7.00	109.61	384.00	95
hypothetical protein, 11.2K	pir2:JC1150	!	7.00	109.14	379.86	102
hypothetical protein HP1511	pir2:G64708	!	7.00	108.76	376.56	108
cytochrome c2 - Agrobacteriu	pir1:CCAG2	+	7.00	108.58	374.99	111
hypothetical 12K protein - E	pir2:S10918	+	7.00	108.52	374.47	112
T-cell receptor beta chain p	pir2:A32578	+	7.00	108.46	373.96	113
hypothetical protein PA0554	pir2:B83575	+	7.00	108.46	373.96	113
hypothetical protein MLCB188	pir2:T44881	+	7.00	108.23	372.47	116
hypothetical protein APE2186	pir2:A82259	-	7.00	108.23	371.98	117
hypothetical protein APE2186	pir2:E72526	+	7.00	108.15	371.02	119
probable DNA binding protein	pir2:T36232	-	7.00	107.95	369.62	122
hypothetical protein APE2026	pir2:D72506	+	7.00	107.85	368.70	124
hypothetical protein PA1969	pir2:A83400	+	7.00	107.53	366.05	130
hypothetical protein APE2472	pir2:H72478	-	7.00	107.48	365.62	131
hypothetical protein F7K2.40	pir2:T05440	+	7.00	107.38	364.78	133
T cell receptor Mb11 beta ch	pir2:S57869	-	7.00	107.19	363.13	137
hypothetical protein ybdD -	pir2:B69747	+	7.00	107.19	363.13	137
hypothetical protein APE0742	pir2:G72664	-	7.00	107.14	362.73	138
hypothetical protein sll1960	pir2:S74898	-	7.00	106.95	361.15	140
T-cell receptor beta chain p	pir2:S38388	-	7.00	106.86	360.38	142
conserved hypothetical prote	pir2:F83544	+	7.00	106.80	359.88	144
conserved hypothetical prote	pir2:B83291	+	7.00	106.68	358.88	148
hypothetical 16.4K protein (	pir2:S56500	+	7.00	106.63	358.51	148
hydrogenase-related protein	pir2:S53659	+	7.00	106.46	357.06	149
trappin-10 - collared peccar	pir2:T29164	+	7.00	106.33	356.06	153
tuberisation-related protein	pir2:JE0256	+	7.00	106.04	353.63	156
hypothetical 17.5 kD protein	pir2:T07174	+	7.00	105.76	351.36	157
hypothetical protein Y - Myx	pir2:G65156	+	7.00	105.72	351.05	163
nonstructural protein C - ca	pir2:A83474	-	7.00	105.61	350.12	170
conserved hypothetical prote	pir2:G83157	+	7.00	105.46	348.91	171
hypothetical protein CT358 -	pir1:MNN2CV	-	7.00	105.38	348.31	174
hypothetical protein DKFZp43	pir2:C71524	-	7.00	105.27	347.43	178
hypothetical protein FII16.9	pir2:T43451	-	7.00	105.13	346.29	180
probable iron-sulfur protein	pir2:T47707	+	7.00	105.03	345.45	183
iron-sulfur protein CooF [si	pir2:S39894	+	7.00	104.96	344.90	187
hypothetical protein APE1760	pir2:B42957	-	7.00	104.89	344.35	190
hypothetical protein AF1001	pir2:F72559	+	7.00	104.89	344.35	192
hypothetical protein A - Bac	pir2:A69375	+	7.00	104.89	344.35	194
hypothetical protein PH1211	pir2:S31841	-	7.00	104.89	344.35	194
hypothetical protein - Deino	pir2:E71064	+	7.00	104.89	344.35	194
ryudocan precursor - mouse	pir2:E75297	-	7.00	104.75	343.28	198
hypothetical protein PA2792	pir2:JC5613	+	7.00	104.75	343.28	198
	pir2:H83297					

us-09-388-090-3.oligo7.rpi

pir id	sign					description
pir2:F70640	-	7.00	104.75	343.28	198	probable transcription regulato
pir2:T64125	+	7.00	104.72	343.02	199	nucleotide-binding protein sfuC
pir2:C82224	-	7.00	104.69	342.76	200	probable alpha-ribazole-5'-phos
pir2:A42410	-	7.00	104.62	342.24	202	ryudocan precursor - rat
pir2:G82029	+	7.00	104.62	342.24	202	probable lipoprotein NMA0339 [i
pir2:A81008	-	7.00	104.59	341.98	203	hemolysin, probable NMB2091 [in
pir2:A45317	+	7.00	104.56	341.72	204	sperm-specific histone H1-like
pir2:T08321	-	7.00	104.43	340.71	208	hypothetical protein H1243 [acc
pir2:T35820	-	7.00	104.30	339.72	212	probable integral membrane prot
pir2:S75787	-	7.00	104.24	339.23	214	methyltransferase - Synechocyst
pir2:D83278	+	7.00	104.06	337.81	220	hypothetical protein PA2941 [im
pir2:T46524	-	7.00	104.06	337.81	220	probable disulfide bond-forming
pir2:D81122	+	7.00	104.00	337.34	222	baseplate assembly protein V, F
pir1:C57987	-	7.00	103.97	337.11	223	nrfC protein - Escherichia coli
pir2:H70717	+	7.00	103.94	336.88	224	hypothetical protein Rv0962c -
pir1:T35058	-	7.00	103.94	336.88	224	hypothetical protein SC4G2.23 S
pir1:A64181	-	7.00	103.91	336.65	225	nrfC protein - Haemophilus infl
pir2:T25532	-	7.00	103.79	335.75	229	conserved hypothetical protein
pir2:E81099	-	7.00	103.76	335.52	230	ribosomal small chain pseudouri
pir2:S22322	-	7.00	103.59	334.21	236	histone H1 - wheat
pir1:D69116	-	7.00	103.57	333.99	237	anaerobic ribonucleotide reduct
pir2:G75316	-	7.00	103.57	333.99	237	probable lipase - Deinococcus r
pir2:T07092	+	7.00	103.51	333.56	239	Ca+2-binding EF hand protein hc
pir2:H72669	+	7.00	103.40	332.72	243	probable glutaredoxin-like prot
pir1:C65048	-	7.00	103.35	332.30	245	hypothetical protein b2682 - Es
pir1:T25533	-	7.00	103.32	332.10	246	hypothetical protein C09B7.1 -
pir2:T16264	-	7.00	103.29	331.89	247	hypothetical protein F2206.180
pir2:E81695	-	7.00	103.27	331.69	248	3-oxoacyl-(acyl carrier protein
pir2:F82943	-	7.00	103.27	331.69	248	probable ABC substrate-binding
pir2:T11218	-	7.00	103.24	331.48	249	probable succinate dehydrogenas
pir2:E75534	-	7.00	103.24	331.48	249	ribosomal protein S3 - Deinococ
pir2:D75317	-	7.00	103.16	331.28	250	hypothetical protein DR2083 - D
pir2:E72478	+	7.00	103.06	330.88	252	probable cobyric acid synthase
pir2:D70376	+	7.00	103.01	330.08	258	gene 38 protein - phage SPP1
pir2:T16264	+	7.00	102.90	329.69	262	probable dehydrogenase - Strept
pir2:F81381	+	7.00	102.90	328.92	262	octopine transport protein occH
pir2:C83275	-	7.00	102.88	328.73	266	hypothetical protein C25B8.4 -
pir2:S25039	-	7.00	102.80	328.16	267	probable dehydrogenase - Mycoba
pir2:S77706	+	7.00	102.78	327.97	271	conserved hypothetical protein
pir2:C83487	+	7.00	102.68	327.23	285	2-dehydro-3-deoxyphosphooctonat
pir2:B43253	-	7.00	102.68	327.23	288	4-amino-4-deoxychorismate lyase
pir2:H81836	-	7.00	102.35	324.72	288	hypothetical protein T21B10.1 -
pir1:C65202	-	7.00	102.28	324.20	291	probable transcription regulato
pir2:T34165	-	7.00	102.28	324.20	292	hypothetical protein YGR239c -
pir2:JC6537	+	7.00	102.21	323.69	297	probable pyruvate formate-lyase
pir2:T51118	-	7.00	102.19	323.52	298	hypothetical protein C42D4.11 -
pir2:JQ0707	-	7.00	102.08	322.68	299	mda-9 protein - human
pir2:T34876	+	7.00	102.05	322.52	300	phytoene synthase crtB [importe
pir1:T35246	+	7.00	102.03	322.35	302	methanol oxidation protein moxJ
pir2:E72463	+	7.00	102.01	322.19	302	probable secreted protein - Str
pir2:B70600	-	7.00	101.96	321.86	303	probable integral membrane prot
pir2:T72456	-	7.00	101.96	321.70	304	hypothetical protein APE2352 -
pir2:H83064	-	7.00	101.94	321.54	305	probable ribokinase - Mycobacte
pir2:G69207	+	7.00	101.92	321.38	305	hydroxymethylbilane synthase [E
pir2:E72670	-	7.00	101.90	321.05	307	conserved hypothetical protein
pir2:S51573	-	7.00	101.90	320.90	308	conserved hypothetical protein
pir2:C70426	-	7.00	101.86	320.74	309	hypothetical protein APE0787 -
pir1:D70503	-	7.00	101.83	320.58	311	mocB protein precursor - Rhizob
pir2:A75514	+	7.00	101.81	320.42	311	RNA polymerase sigma factor Rpo
pir2:E69822	+	7.00	101.77	320.26	312	probable integrase/recombinase
pir2:A75258	+	7.00	101.75	320.11	313	peptidase-related protein - Dei
pir2:B82334	+	7.00	101.71	319.95	314	hypothetical protein yhcI - Bac
pir2:T35857	-	7.00	101.69	319.80	315	conserved hypothetical protein
pir2:G82752	+	7.00	101.58	319.03	320	tRNA delta(2)-isopentenylpyroph
pir2:T50875	+	7.00	101.56	321.05	320	cobalt-zinc-cadmium resistance
pir2:A44504	+	7.00	101.52	318.88	321	hypothetical electron transfer
pir2:T82435	-	7.00	101.52	318.57	323	suhR protein - Rhizobium melilo
pir2:A81213	+	7.00	101.50	318.42	324	D-3-phosphoglycerate dehydrogen
						octaprenyl-diphosphate synthase

description		pir id	sign				
36.3K zipper-containing prot	324	pir2:S43424	+	7.00	101.50	318.42	
necdin, brain - mouse	325	pir2:JN0148	+	7.00	101.48	318.27	
probable secreted serine-ric	327	pir2:T36120	-	7.00	101.44	317.83	
probable electron transfer 1	328	pir2:H82971	+	7.00	101.42	317.83	
hypothetical protein F54C9.1	328	pir2:T22636	-	7.00	101.42	317.68	
hypothetical protein yddH [i	329	pir2:A69776	+	7.00	101.40	317.10	
iron(III) dicitrate transpor	333	pir2:D69805	+	7.00	101.32	316.81	
probable vitamin B12 ABC tra	335	pir2:G82224	-	7.00	101.28	316.81	
replication-associated prote	335	pir1:A42654	-	7.00	101.28	316.81	
alcohol dehydrogenase (EC 1.	337	pir1:DEBHG	-	7.00	101.24	316.52	
glyceraldehyde-3-phosphate d	337	pir2:T06720	-	7.00	101.24	316.52	
hypothetical protein F28P10.	337	pir1:S45605	-	7.00	101.24	316.52	
alcohol dehydrogenase (EC 1.	339	pir2:F70625	+	7.00	101.20	316.23	
hypothetical protein Rv1043c	341	pir2:C83578	-	7.00	101.16	315.95	
hypothetical protein PA0536	341	pir2:G83127	-	7.00	101.16	315.95	
conserved hypothetical prote	346	pir2:B59550	-	7.00	101.07	315.25	
hypothetical protein AF2401	348	pir2:E82307	-	7.00	101.03	314.97	
proteinase DegS VC0565 [impo	352	pir2:H70636	+	7.00	100.95	314.42	
probable riba protein - Myco	353	pir2:T37317	-	7.00	100.93	314.29	
probable Ca2+/calmodulin-dep	357	pir2:T36783	-	7.00	100.86	313.75	
probable zinc-binding alcoho	358	pir2:G70814	-	7.00	100.84	313.61	
probable far protein - Mycob	359	pir2:T06197	-	7.00	100.82	313.48	
HC-toxin reductase (EC 1...	362	pir2:S77203	-	7.00	100.77	313.08	
hypothetical protein slr1821	366	pir2:A82094	-	7.00	100.69	312.56	
thiamin biosynthesis lipopro	367	pir2:B45610	-	7.00	100.68	312.43	
aldolase ALDO-1 - Plasmodium	368	pir1:PABY21	-	7.00	100.66	312.30	
phosphoprotein phosphatase (	369	pir1:TVFVAF	-	7.00	100.64	312.17	
transforming protein maf - a	369	pir2:A44942	-	7.00	100.64	312.17	
fructose-bisphosphate aldola	369	pir2:G02131	+	7.00	100.64	312.17	
proteinase-activated recepto	369	pir2:A75371	-	7.00	100.64	312.17	
threonine 3-dehydrogenase -	370	pir2:C69309	+	7.00	100.62	312.04	
conserved hypothetical prote	370	pir2:H64545	+	7.00	100.62	312.04	
lipopolysaccharide 1,2-gluco	370	pir2:S04880	+	7.00	100.62	312.04	
probable transposase I (inse	371	pir2:D69253	+	7.00	100.61	311.91	
conserved hypothetical prote	371	pir2:H83285	-	7.00	100.61	311.91	
probable transposase T31 - E	372	pir1:QQEC47	+	7.00	100.59	311.78	
hypothetical protein MLCB637	372	pir2:T45410	-	7.00	100.59	311.78	
probable membrane protein YD	372	pir2:S77770	-	7.00	100.59	311.78	
probable lpp2 protein - Myco	373	pir2:A70856	+	7.00	100.57	311.65	
recF protein - Haemophilus i	375	pir2:G64482	-	7.00	100.53	311.40	
probable GTP-binding pro	377	pir1:I64106	-	7.00	100.46	311.15	
phosphoprotein phosphatase (	379	pir1:PABY22	+	7.00	100.41	310.89	
hypothetical protein RO5C11.	382	pir2:T33875	-	7.00	100.41	310.52	
probable Cynx-related transp	382	pir2:C83337	-	7.00	100.38	310.52	
hypothetical protein PA2465	384	pir2:D75541	-	7.00	100.36	310.27	
probable serine proteinase -	385	pir2:T28691	+	7.00	100.36	310.15	
isochorismate synthase (EC 5	385	pir1:SYECIK	-	7.00	100.26	310.15	
replication protein L28 - Al	391	pir2:S60672	-	7.00	100.26	309.42	
probable MFS transporter PA3	391	pir2:F83233	-	7.00	100.26	309.42	
3-methyl-2-oxobutanoate dehy	391	pir1:A34267	+	7.00	100.24	309.30	
dom-3 protein - Caenorhabdit	392	pir2:S60465	+	7.00	100.22	309.18	
4-hydroxybenzoate 3-monooxyg	393	pir1:WHPSBF	+	7.00	100.21	309.06	
4-hydroxybenzoate 3-monooxyg	394	pir1:WHPSBA	-	7.00	100.21	309.06	
probable membrane protein YP	394	pir2:S62009	+	7.00	100.21	309.06	
p-hydroxybenzoate hydroxylas	394	pir2:B83613	-	7.00	100.21	308.82	
proteinase-activated recepto	396	pir2:T02483	-	7.00	100.17	308.70	
proteinase-activated recepto	397	pir2:S66518	-	7.00	100.16	308.70	
probable transaminase (EC 2.	397	pir2:S64709	-	7.00	100.16	308.70	
proteinase activated recepto	397	pir2:B70815	-	7.00	100.16	308.46	
coat protein precursor - bla	399	pir1:I48705	-	7.00	100.12	307.53	
hypothetical protein F2809.4	407	pir1:VCBB2G	-	7.00	99.99	307.30	
probable carboxyl-terminal p	409	pir2:T45797	-	7.00	99.96	307.07	
hypothetical protein 23 - eq	411	pir2:S26195	-	7.00	99.93	306.62	
conserved hypothetical prote	415	pir2:S55617	+	7.00	99.86	306.62	
glycoprotein qD - bovine her	415	pir2:C75592	-	7.00	99.86	306.39	
hypothetical protein BO393.6	417	pir1:VGBEIB	-	7.00	99.83	306.39	
probable hydrolase - Deinoco	417	pir2:S35784	-	7.00	99.83	306.39	
	417	pir2:T18740	-	7.00	99.83	306.28	
	418	pir2:F75472	-	7.00	99.82		

us-09-388-090-3.oligo7.rpr

Right-hand table

Description	No.	Score	%Id	7.00	+/-	PIR ID
cytochrome P450 4F6 protein	537	294.79	98.16	7.00	-	pir2:JC4534
capsid protein - human cali	539	294.62	98.14	7.00	+	pir2:S40111
probable integral membrane	542	294.37	98.10	7.00	-	pir2:C70732
alginate-c5-mannuronan-epim	543	294.29	98.06	7.00	+	pir2:F83202
H+-transporting ATP synthas	545	294.13	98.05	7.00	+	pir1:PWBYA
DNA-binding protein pabf - (	546	294.04	98.03	7.00	+	pir2:T02029
hypothetical protein C16B8.	548	293.88	98.02	7.00	+	pir2:T15523
probable htrA protein - Myc	549	293.80	98.02	7.00	+	pir2:B70508
ABC transporter integral mem	553	293.47	97.97	7.00	-	pir2:T35451
hemolysin accessory fhaC pr	554	293.39	97.96	7.00	+	pir2:S41526
long-chain-fatty-acid--CoA l	562	292.75	97.86	7.00	-	pir2:C83233
cellulase (EC 3.2.1.4) B pr	563	292.67	97.85	7.00	+	pir1:CZCLBM
probable deaD protein - Myc	563	292.67	97.81	7.00	+	pir2:E70752
leucyl aminopeptidase (EC 3	566	292.43	97.81	7.00	+	pir2:T07850
protein disulfide-isomerase	566	292.43	97.80	7.00	+	pir2:T06724
hypothetical protein SPAC6F1	567	292.36	97.78	7.00	+	pir2:T11653
leucyl aminopeptidase (EC 3	569	292.20	97.78	7.00	+	pir2:S57812
urease (EC 3.5.1.5) alpha ch	569	292.20	97.76	7.00	+	pir2:S75169
C conserved hypothetical pro	569	292.20	97.73	7.00	+	pir2:A69511
leucyl aminopeptidase (EC 3	571	292.04	97.57	7.00	+	pir2:T07849
leucyl aminopeptidase (EC 3	571	292.04	97.57	7.00	+	pir2:T07047
leucyl aminopeptidase (EC 3	573	291.89	97.56	7.00	+	pir1:S41376
beta-fructofuranosidase (EC	587	290.81	97.56	7.00	+	pir2:T08439
conserved hypothetical prote	587	290.81	97.54	7.00	+	pir2:T08829
methyl-accepting chemotaxis	587	290.81	97.52	7.00	+	pir2:E82431
hypothetical protein SC6G3.0	588	290.74	97.48	7.00	+	pir2:T35549
hypothetical protein - fruit	588	290.74	97.47	7.00	-	pir2:T31675
hypothetical protein SC7A1.0	589	290.66	97.37	7.00	+	pir2:T35649
hypothetical protein C07A9.1	590	290.59	97.24	7.00	-	pir2:S40707
hypothetical protein 22E5.7	592	290.44	97.19	7.00	-	pir2:T13742
hypothetical protein PAB1294	595	290.21	97.17	7.00	+	pir2:F75008
succinate dehydrogenase, fla	596	290.14	97.05	7.00	+	pir2:E82728
probable membrane protein YG	605	289.48	97.05	7.00	+	pir2:S64410
probable membrane protein YG	614	288.82	96.96	7.00	+	pir2:S64447
dihydroxy-acid dehydratase -	617	288.61	96.93	7.00	+	pir2:T36294
probable membrane protein YP	622	288.25	96.93	7.00	+	pir2:S61140
hypothetical protein PA2044	624	288.11	96.91	7.00	+	pir2:C83389
involucrin L - gorilla	635	287.34	96.91	7.00	-	pir1:I37060
collagen alpha 2(VIII) chain	635	287.34	96.90	7.00	-	pir2:A57131
hypothetical protein B23t21.	644	286.73	96.89	7.00	+	pir2:T49694
epidermal growth factor rece	647	286.53	96.87	7.00	+	pir2:A36325
penicillin-binding protein h	647	286.53	96.78	7.00	+	pir2:F70057
probable membrane protein YI	648	286.46	96.77	7.00	+	pir2:S48471
probable serine-threonine-pr	649	286.39	96.69	7.00	+	pir2:T37581
PTS system, mannitol-specifi	649	286.39	96.61	7.00	-	pir2:B82385
hypothetical protein F45H7.3	650	286.32	96.38	7.00	-	pir2:T22254
Na+/Ca2+,K+-exchanging prote	651	286.26	96.37	7.00	+	pir2:S44806
penicillin-binding protein C	653	286.12	96.22	7.00	+	pir2:T03889
receptor-like protein kinase	662	285.52	96.11	7.00	+	pir2:E72080
glutamine transport protein,	663	285.46	96.05	7.00	+	pir2:T46005
Om(2D) protein - fruit fly (	671	284.94	95.92	7.00	-	pir2:S65566
regulatory protein flaY - Ca	679	284.42	95.88	7.00	-	pir2:S51599
nodulation protein-related p	703	282.92	95.79	7.00	+	pir2:S02165
Na+/Ca2+,K+-exchanging prote	703	282.92	95.45	7.00	+	pir2:G75638
microsomal endopeptidase MEP	704	282.86	95.42	7.00	-	pir2:T03888
transponson protein MEP	710	282.49	95.23	7.00	-	pir2:A45985
probable membrane protein yc	720	281.89	95.16	7.00	+	pir1:T31197
gene su(f) protein, 84K spli	733	281.12	95.15	7.00	+	pir2:G64836
primosomal replication facto	734	281.06	95.15	7.00	+	pir1:A64389
sensory transduction histidi	739	280.77	95.11	7.00	-	pir2:A82047
hypothetical protein F1BE5.1	749	280.19		7.00		pir2:T05163
ferrichrome receptor precurs	754	279.91		7.00		pir2:S77175
hypothetical protein F4P12.4	758	279.68		7.00		pir2:T45910
hypothetical protein F45H7.2	769	279.07		7.00		pir2:S30948
hypothetical protein wrm-1	796	277.60		7.00		pir2:T22256
probable leukotriene-A4 hydr	809	276.92		7.00		pir2:T03704
hypothetical protein flay - Ca	813	276.71		7.00		pir2:T32899
glycoprotein H precursor - b	842	275.23		7.00		pir2:S76607
hypothetical protein F28J12.	846	275.03		7.00		pir2:S18462
serine /threonine kinase (EC	847	274.98		7.00		pir2:T04533
receptor-like kinase (EC 2.7	847	274.98		7.00		pir2:T14445
probable membrane protein YD	852	274.74		7.00		pir2:S57984

Left-hand table

PIR ID	No.	Description	Score	%Id	7.00	+/-
pir2:G75020	419	glucose-1-phosphate thymidylyl	306.17	99.80	7.00	+
pir2:S42289	419	T48 protein - fruit fly (Drosop	306.17	99.80	7.00	+
pir2:F83494	420	hypothetical protein PA1220 [im	306.06	99.74	7.00	+
pir2:B70178	423	vacuolar X-prolyl dipeptidyl am	305.73	99.71	7.00	+
pir2:E71982	425	isocitrate dehydrogenase - Heli	305.51	99.69	7.00	+
pir2:C64523	425	isocitrate dehydrogenase (NADP+	305.51	99.66	7.00	+
pir2:S04848	426	transcription initiation factor	305.40	99.46	7.00	-
pir2:D75582	428	hypothetical protein - Deinococ	305.18	99.37	7.00	+
pir2:E82583	430	threonine synthase XF2223 (impc	304.96	99.33	7.00	-
pir2:T36544	441	probable two-component system s	303.79	99.33	7.00	+
pir1:T31539	446	hypothetical protein Y47D3A.7 -	303.27	99.32	7.00	+
pir2:T35519	447	hypothetical protein SC6G10.10c	303.16	99.26	7.00	-
pir1:F64637	450	D-alanine/glycine transport pro	302.85	99.24	7.00	+
pir2:H64741	450	yaeL protein - Escherichia coli	302.85	99.24	7.00	-
pir2:T47824	451	hypothetical protein F24G16.270	302.75	99.21	7.00	+
pir2:JC6051	455	trypsin-like proteinase (EC 3.4	302.34	99.20	7.00	+
pir2:F82307	456	proteinase DO VC0566 [imported]	302.24	99.18	7.00	-
pir2:T21822	457	hypothetical protein F35H8.5 -	302.24	99.18	7.00	+
pir2:B39200	458	siroheme synthase - Salmonella	302.14	99.11	7.00	-
pir2:T27745	459	probable pyruvate kinase APE04B	302.04	99.11	7.00	+
pir2:G71641	459	dihydrolipoamide dehydrogenase	301.94	99.06	7.00	+
pir2:T36334	460	hypothetical protein SCE9.15c	301.94	99.06	7.00	-
pir2:T10995	460	NADH dehydrogenase (ubiquinone)	301.84	99.06	7.00	+
pir2:I76771	465	phosphomannomutase (EC 5.4.2.8)	301.84	99.06	7.00	-
pir2:T34936	466	hypothetical protein SC3F9.14 S	301.34	99.01	7.00	+
pir2:S76657	469	coproporphyrinogen oxidase (EC	301.24	98.97	7.00	-
pir1:A48788	469	leucyl aminopeptidase (EC 3.4.1	300.95	98.96	7.00	+
pir2:A24450	469	collagen alpha 2(VIII) chain -	300.95	98.93	7.00	-
pir2:T35526	469	probable integral membrane tran	300.95	98.92	7.00	+
pir2:B83195	469	probable MFS transporter PA3595	300.95	98.89	7.00	+
pir2:D70542	472	hypothetical protein Rv1586c -	300.65	98.88	7.00	-
pir2:C70853	475	hypothetical protein Rv3087 - M	300.36	98.88	7.00	-
pir2:T23793	476	hypothetical protein M18.3 - Ca	300.27	98.86	7.00	+
pir2:A70318	478	aldehyde dehydrogenase - Aquife	300.08	98.84	7.00	+
pir2:T43042	479	hypothetical protein - fission	299.98	98.82	7.00	-
pir2:A71273	481	hypothetical protein TP0865 - s	299.79	98.82	7.00	+
pir1:T43809	482	methylaspartate mutase (EC 5.4	299.69	98.79	7.00	+
pir2:G75483	482	probable leucyl aminopeptidase	299.69	98.79	7.00	+
pir2:S22654	483	ARS-binding factor-1 - yeast (K	299.60	98.70	7.00	+
pir2:H70112	485	periplasmic serine proteinase D	299.41	98.70	7.00	+
pir2:S75684	486	hypothetical protein F08D12.12	299.32	98.66	7.00	+
pir2:T50749	486	methoxyneurosporene dehydrogena	299.32	98.63	7.00	+
pir2:S49624	488	methoxyneurosporene dehydrogena	299.13	98.57	7.00	+
pir2:D70876	488	probable polyketidesynthase bet	299.13	98.57	7.00	+
pir2:A82984	495	hypothetical protein PA5294 [im	298.48	98.50	7.00	-
pir1:I37062	495	involucrin S - gorilla	298.48	98.49	7.00	-
pir2:S23633	495	methoxyneurosporene dehydrogena	298.20	98.46	7.00	+
pir2:T46176	498	probable cytochrome P450 T4D2.2	298.02	98.46	7.00	+
pir2:F70012	500	phosphoglycerate kinase homolog	297.57	98.46	7.00	-
pir1:KICRGF	505	receptor-like protein kinase	297.57	98.39	7.00	-
pir2:T46996	505	hypothetical protein aruD [impo	297.48	98.39	7.00	+
pir2:A56391	506	lamina associated polypeptide 1	297.12	98.34	7.00	-
pir2:B70368	510	L-aspartate oxidase - Aquifex a	297.03	98.34	7.00	+
pir2:S10527	511	endoglucanase B precursor - Pse	296.86	98.31	7.00	+
pir2:H71722	513	probable periplasmic serine pro	296.86	98.29	7.00	+
pir2:T40998	513	hypothetical protein SPCC1450.1	296.77	98.29	7.00	+
pir2:A82432	514	sodium/solute symporter VCA0667	296.33	98.27	7.00	-
pir2:B83360	519	hypothetical protein PA2286 [im	296.33	98.26	7.00	+
pir2:A82634	519	conserved hypothetical protein	296.33	98.25	7.00	+
pir2:H69034	519	leucyl aminopeptidase (EC 3.4.1	295.98	98.21	7.00	+
pir2:D70695	523	hydroxymethylpyrimidine kinase	295.98	98.19	7.00	-
pir2:S57811	525	pituitary adenylate cyclase act	295.81	98.17	7.00	-
pir2:T07834	527	hypothetical protein F27K19.180	295.64		7.00	+
pir2:JN0902	527	hypothetical protein ZC434.9 -	295.64		7.00	
pir2:T49212	528	L1 protein - human papillomavir	295.55		7.00	
pir2:T27572	528	hypothetical 60.8 kD protein in	295.47		7.00	
pir2:S36520	529	55.5K sporulation protein - Str	295.38		7.00	
pir2:D65214	530	probable serine rich protein	295.13		7.00	
pir2:JQ0783	533	probable serine proteinase htrA	294.96		7.00	
pir2:S62439	533	F54F2.2 protein - Caenorhabditi	294.88		7.00	
pir2:T45197	535	noncapsid protein NS1 - adeno-a				
pir2:S44827	536					
pir1:UYADIA						+

						ID	Description
+	7.00	95.05	274.35	860	pir2:S43846	xylanase B - rumen fungus (Neoc
+	7.00	94.94	273.67	874	pir2:T15570	hypothetical protein C23F12.2
+	7.00	94.91	273.48	878	pir2:T15621	hypothetical protein F32A7.5 -
+	7.00	94.89	273.34	881	pir2:F72397	pyruvate,orthophosphate dikinas
-	7.00	94.88	273.29	882	pir2:PC1231	hypothetical protein 882 (3' re
+	7.00	94.86	273.15	885	pir2:T18387	hypothetical protein SPAC25H1.0
+	7.00	94.85	273.10	886	pir2:S54355	inter-alpha-trypsin inhibitor h
-	7.00	94.84	273.01	888	pir2:T01081	hypothetical protein T10P11.3.
+	7.00	94.71	272.22	905	pir2:T23229	hypothetical protein ZC101.1 -
+	7.00	94.69	272.08	908	pir2:C70168	DNA polymerase I (polA) homolog
+	7.00	94.63	271.72	916	pir2:G75417	SNF2/Rad54 helicase-related pro
+	7.00	94.59	271.45	922	pir1:S54342	protein-tyrosine-phosphatase (E
-	7.00	94.59	271.45	922	pir2:T37256	metalloproteinase sup-17 - Caen
+	7.00	94.47	270.69	939	pir2:S28394	probable serine/threonine-speci
-	7.00	94.42	270.38	946	pir2:S27921	nuclear antigen EBNA-3B - human
+	7.00	94.30	269.61	964	pir2:S15342	hypothetical protein B0336.1 -
+	7.00	94.27	269.44	968	pir2:T00353	hypothetical protein KIAA0697
+	7.00	94.19	268.97	979	pir1:JC2349	protein-tyrosine-phosphatase (E
+	7.00	94.08	268.27	996	pir2:I48721	PTP 35 protein - mouse
+	7.00	94.03	267.98	1003	pir2:T28654	transposase - Pseudomonas puti
+	7.00	94.01	267.82	1007	pir2:PN0156	glutamate receptor channel del
+	7.00	94.00	267.78	1008	pir2:S28858	glutamate receptor delta-2 cha
-	7.00	93.92	267.29	1021	pir2:T42229	probable E1-E2 ATPase (EC 3.6.
+	7.00	93.92	267.25	1021	pir2:S44644	hypothetical protein F37A4.7
+	7.00	93.83	266.70	1035	pir2:T42093	phospholipase D (EC 3.1.4.4) 1
+	7.00	93.82	266.66	1035	pir2:T13732	phospholipase D (EC 3.1.4.4) 1
+	7.00	93.82	266.66	1036	pir2:T18530	phospholipase D (EC 3.1.4.4) 1
+	7.00	93.82	266.66	1036	pir2:T34258	phospholipase D (EC 3.1.4.4) 1
+	7.00	93.81	266.62	1037	pir2:T13943	phospholipase D (EC 3.1.4.4) 1
-	7.00	93.80	266.50	1040	pir2:T29092	TSC-22 protein homolog - chick
+	7.00	93.74	266.19	1048	pir1:YGVCAR	angR protein - Vibrio anguilla
+	7.00	93.74	266.19	1048	pir2:T23764	hypothetical protein M142.6 -
+	7.00	93.74	266.15	1049	pir1:A27079	fibronectin receptor alpha cha
+	7.00	93.71	266.00	1053	pir2:S44250	integrin alpha-5 chain - mouse
+	7.00	93.61	265.39	1069	pir2:D81826	exodeoxyribonuclease V (EC 3.1
+	7.00	93.61	265.39	1069	pir2:A81050	exodeoxyribonuclease V (125 kD
+	7.00	93.61	265.39	1069	pir2:T42681	hypothetical protein DKFZp434E
+	7.00	93.58	265.20	1074	pir2:T13725	phospholipase (EC 3.1.4.-) D1
+	7.00	93.58	265.16	1075	pir2:T46635	phospholipase D (EC 3.1.4.4) 1
+	7.00	93.56	265.08	1077	pir2:T20579	phospholipase D (EC 3.1.4.4) 1
-	7.00	93.49	264.64	1089	pir2:T14576	nosA protein - slime mold (Dic
-	7.00	93.45	264.41	1095	pir2:T43275	neurabin - rat
+	7.00	93.40	264.08	1104	pir2:C72409	reverse gyrase - Thermotoga ma
+	7.00	93.36	263.87	1110	pir2:A43253	Large tra-1 protein - Caenorha
-	7.00	93.32	263.61	1117	pir2:S63399	probable membrane protein YNR0
-	7.00	93.30	263.47	1121	pir2:S54504	hypothetical protein YPR030w
+	7.00	93.28	263.33	1125	pir2:B41206	microtubule-associated protein
+	7.00	93.24	263.08	1132	pir2:T49403	related to protein-tyrosine-ph
-	7.00	93.23	263.04	1133	pir2:S54498	probable membrane protein YPRO
+	7.00	93.12	262.37	1152	pir2:S20106	hypothetical protein 2 - slime
+	7.00	93.12	262.37	1152	pir2:A33183	microtubule-associated protein
+	7.00	93.00	261.65	1173	pir2:I50620	prockr2 - chicken (fragment)
+	7.00	92.97	261.45	1179	pir2:T04584	TMV resistance protein N homol
-	7.00	92.90	261.08	1190	pir2:S47536	SWH1 protein - yeast (Saccharo
+	7.00	92.90	261.04	1191	pir2:A53491	bumetanide-sensitive Na-K-Cl (
+	7.00	92.88	260.91	1195	pir1:S26722	DNA-directed RNA polymerase (E
+	7.00	92.82	260.55	1206	pir2:S72620	probable reverse transcriptase
+	7.00	92.78	260.35	1212	pir2:T00332	hypothetical protein KIAA0559
+	7.00	92.76	260.19	1217	pir2:S52714	sericinlB - silkworm
+	7.00	92.75	260.12	1219	pir2:S54570	probable membrane protein YMRI
+	7.00	92.66	259.60	1235	pir1:QQBEW4	DNA-binding protein - human cy
+	7.00	92.64	259.51	1238	pir2:T34929	hypothetical protein SC3F9.07
+	7.00	92.62	259.38	1242	pir2:T20739	hypothetical protein F11A10.1
+	7.00	92.54	258.87	1258	pir2:T29041	hypothetical protein B0228.4
-	7.00	92.47	258.47	1271	pir2:S37958	myosin heavy chain homolog YK1
+	7.00	92.10	256.24	1345	pir2:T00964	hypothetical protein F20D22.14
+	7.00	92.06	256.01	1353	pir1:JH0675	restrictin precursor - chicken
-	7.00	91.96	255.41	1374	pir1:VCBE17	major capsid protein - human h
+	7.00	91.84	254.74	1398	pir2:S56814	microtubule-interacting protei
+	7.00	91.73	254.08	1422	pir2:T42636	protein-tyrosine-phosphatase (
+	7.00	91.71	253.99	1425	pir2:T31153	hypothetical protein 235 - Sp1
+	7.00	91.64	253.53	1442	pir1:B48148	protein-tyrosine-phosphatase (

ID						Description
pir1:A48148	1445	253.45	91.62	7.00	+	protein-tyrosine-phosphatas
pir2:T13709	1454	253.21	91.58	7.00	+	diacylglycerol kinase (EC 2
pir2:A44345	1468	252.84	91.52	7.00	-	nucleoporin - rat
pir2:D41214	1557	250.58	91.13	7.00	+	protein-tyrosine-phosphatas
pir2:C41214	1630	248.84	90.83	7.00	-	protein-tyrosine-phosphatas
pir1:CGHU6B	1691	248.45	90.58	7.00	+	collagen alpha 6(IV) chain
pir2:A33988	1692	247.42	90.58	7.00	+	adenylate cyclase (EC 4.6.1
pir1:I38902	1721	246.78	90.47	7.00	-	retinoblastoma binding prot
pir2:A56602	1734	246.50	90.42	7.00	+	microtubule-associated ser1
pir2:S37827	1764	242.85	90.30	7.00	+	hypothetical protein YKL014
pir2:T00637	1922	242.66	89.74	7.00	+	hypothetical protein H_GS54
pir2:T00093	2055	240.19	89.30	7.00	+	hypothetical protein KIAA04
pir2:S72705	2118	239.09	89.10	7.00	+	mycocerosate synthase (EC 2
pir2:A44467	2139	238.73	89.03	7.00	+	voltage-dependent calcium c
pir2:JH0426	2140	238.71	89.03	7.00	-	voltage-dependent calcium c
pir2:JH0427	2143	238.66	89.02	7.00	-	voltage-dependent calcium c
pir2:S11339	2166	238.27	88.95	7.00	+	calcium channel protein - r
pir1:JC2349	2171	238.19	88.93	7.00	+	calcium channel alpha-1 cha
pir2:T20145	2172	238.17	88.93	7.00	+	hypothetical protein C52A11
pir2:A45290	2220	237.38	88.79	7.00	-	calcium channel protein typ
pir2:T31733	2327	235.68	88.47	7.00	+	hypothetical protein T05C3.
pir2:T34022	2476	233.45	88.06	7.00	-	zonadhesin - pig
pir2:T16743	2531	232.67	87.92	7.00	+	hypothetical protein R13F6.
pir2:B48666	2897	227.93	87.03	7.00	-	cell proliferation antigen
pir2:S76109	3029	226.38	86.73	7.00	+	hypothetical protein - Syne
pir1:A48666	3256	223.90	86.25	7.00	-	cell proliferation antigen
pir1:A42996	3341	223.02	86.08	7.00	+	genome polyprotein - mosqui
pir2:S43048	3519	221.26	85.74	7.00	+	polyketide synthase type I
pir2:B83641	3535	221.11	85.71	7.00	+	probable hemagglutinin PA00
pir2:T23433	3672	219.83	85.46	7.00	+	hypothetical protein K08C7.
pir2:T37316	3704	219.54	85.40	7.00	+	probable laminin alpha chai
pir1:S02392	4544	212.79	84.05	7.00	-	alpha-2-macroglobulin recep
pir1:S25111	4545	212.78	84.05	7.00	-	alpha-2-macroglobulin recep
pir2:T17464	5069	209.27	83.33	7.00	+	rifamycin polyketide syntha
pir2:T44807	5369	205.96	82.95	7.00	+	mycosubtilin synthetase cha
pir2:C83339	5627	205.96	82.64	7.00	+	hypothetical protein PA2462
pir2:T14593	9376	189.27	79.27	7.00	-	syringomycin synthetase - P

seq_name: pir2:B81914

seq_documentation_block:
probable periplasmic serine proteinase (EC 3.4.21.-) NMA0710 [imported] - Neisseria m
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 19-May-2000
C:Accession: B81914
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
R:Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: A81775; MUID:20222556
A:Accession: B81914
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-499 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB83996.1; PID:g737
A:Experimental source: serogroup A. strain Z2491
C:Genetics:
A:Gene: NMA0710
C:Superfamily: Helicobacter serine proteinase
C:Keywords: hydrolase: serine proteinase

alignment_scores:
    Quality: 127.00   Length: 127
     Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x B81914

Align seg 1/1   to: B81914   from: 1   to: 499

583 TTGAAACCGGGCGAATGGTGCGTCGCTGCCATCGGCGCGCGCCCTTCGGCTTTGA 632

```
      ||||||||||||||||||||||||||||||
      ||||||||||||||||||||||||||||||
195 LeuLysProGlyGluTrpValAlaAlaIleGlyAlaProPheGlyPheAs 211
633 CAACAGCGTGACCGCGCATCGTCCGCCAAAGGCAGAAGCCTGCCA 682
211 pAsnSerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProA 228
683 ACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCCATCAATCCGGGC 732
228 snGluSerTyrThrProPheIleGlnThrAspValAlaIleAsnProGly 244
733 AATTCCGGCGGCCGTGTTCAACTTAAAAGGACAGGTCGTCCGCATCAA 782
245 AsnSerGlyGlyProLeuPheAsnLeuLysGlySerGlnValValGlyIleIleAs 261
783 TTCGCAAATATACAGCCGGATTCATGGCCATCTCCTTTGCCA 832
261 nSerGlnIleTyrSerArgSerGlyGlyPheMetGlyIleSerPheAlaI 278
833 TCCCGATTGACGTTGCCATGAATGTCGCCGAACAGCGTGAAAAACACCGGC 882
278 leProIleAspValAlaMetAsnValAlaGluGlnLeuLysAsnThrGly 294
883 AAAGTCCAACCGGGACAACTGGGCGTGATTATTCAGGAAGTATCCTACGG 932
295 LysValGlnArgGlyGlyGlnLeuGlyValIleIleGlnGluValSerTyrGl 311
933 TTTGGCACAGTCGTTCGGTCTGGATAAAGCC 963
311 yLeuAlaGlnSerPheGlyLeuAspLysAla 321
```

seq_name: pir2:F83550

seq_documentation_block:
serine proteinase MucD precursor PA0766 [imported] - Pseudomonas aeruginosa (strain PAO1
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C;Accession: F83550
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950
A;Accession: F83550
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <STO>
A;Cross-references: GB:AE004511; GB:AE004091; NID:g9946646; PIDN:AAG04155.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: mucD; PA0766

alignment_scores:
　　　　Quality:　19.00　　　　Length:　19
　　　　　Ratio:　1.000　　　　　Gaps:　　0
Percent Similarity: 100.000　　Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x F83550

Align seq 1/1 to: F83550 from: 1 to: 474

637 AGGGTGACCGCGCATCGTCCGCCAAAGGCAGAAGCCTGCCACCGA 686
     ||||||||||||||||||||||||||||||||||||||||||||
184 SerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProAsnGl 200

687 AAGCTAC 693
     |||||||
200 uSerTyr 202

seq_name: pir2:F70048

seq_documentation_block:
serine proteinase Do homolog yvtB - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: F70048
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga
liech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,
Koeter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scan
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A;Accession: F70048
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-224 <KUN>
A;Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15289.1; PID:e1184
A;Experimental source: strain 168
C;Genetics:
A;Gene: yvtB

alignment_scores:
　　　　Quality:　11.00　　　　Length:　11
　　　　　Ratio:　1.000　　　　　Gaps:　　0
Percent Similarity: 100.000　　Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x F70048

Align seq 1/1 to: F70048 from: 1 to: 224

718 GCCATCAATCCGGGCAATTCCGGCGGCCGCCGCTG 750
     ||||||||||||||||||||||||||||||||||
58 AlaIleAsnProGlyAsnSerGlyGlyProLeu 68

seq_name: pir2:S74643

seq_documentation_block:
proteinase hhoA (EC 3.4.-.-) - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein sll1679
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S74643
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A;Reference number: S74322; MUID:97061201
A;Accession: S74643
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-394 <KAN>
A;Cross-references: EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BAA16795.1; PID:g165
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: hhoA
C;Superfamily: proteinase hhoB
C;Keywords: hydrolase; proteinase

```
alignment_scores:
         Quality:   11.00             Length:      11
          Ratio:    1.000             Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x S74643

Align seg 1/1  to: S74643  from: 1  to: 394

   718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
       |||||||||||||||||||||||||||||||||||
   231 AlaIleAsnProGlyAsnSerGlyGlyProLeu 241

seq_name: pir2:B71284
```

```
seq_documentation_block:
probable periplasmic serine proteinase DO (htrA-1) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 24-Nov-1999
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwir
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770
A;Accession: B71284
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-398 <COL>
A;Cross-references: GB:AE001248; GB:AE000520; NID:g3323074; PID:g3323308
C;Genetics:
C;Gene: TP0773
C;Superfamily: proteinase hhoB
```

```
alignment_scores:
         Quality:   11.00             Length:      11
          Ratio:    1.000             Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x B71284

Align seg 1/1  to: B71284  from: 1  to: 398

   718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
       |||||||||||||||||||||||||||||||||||
   223 AlaIleAsnProGlyAsnSerGlyGlyProLeu 233

seq_name: pir2:S75445
```

```
seq_documentation_block:
proteinase hhoB (EC 3.4.-.-) - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein sll1427
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201
A;Accession: S75445
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-416 <KAND>
A;Cross-references: EMBL:D90911; GB:AB001339; NID:g1653083; PIDN:BAA18006.1; PID:g165309
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: hhoB
C;Superfamily: proteinase hhoB
C;Keywords: hydrolase; proteinase
```

```
alignment_scores:
         Quality:   11.00             Length:      11
          Ratio:    1.000             Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x S75445

Align seg 1/1  to: S75445  from: 1  to: 416

   718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
       |||||||||||||||||||||||||||||||||||
   252 AlaIleAsnProGlyAsnSerGlyGlyProLeu 262

seq_name: pir2:E75357
```

```
seq_documentation_block:
probable periplasmic serine proteinase Do - Deinococcus radiodurans
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: E75357
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C...
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896
A;Accession: E75357
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <WHI>
A;Cross-references: GB:AE002017; GB:AE000513; NID:g6459527; PIDN:AAF11312.1; PID:g645
A;Experimental source: strain R1
C;Genetics:
C;Gene: DR1756
A;Map position: 1
C;Superfamily: proteinase hhoB
```

```
alignment_scores:
         Quality:   11.00             Length:      11
          Ratio:    1.000             Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x E75357

Align seg 1/1  to: E75357  from: 1  to: 441

   718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
       |||||||||||||||||||||||||||||||||||
   256 AlaIleAsnProGlyAsnSerGlyGlyProLeu 266

seq_name: pir2:A69643
```

```
seq_documentation_block:
serine proteinase DO, heat-shock protein htrA - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: A69643
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fujita, M.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, t
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Larding
```

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033
A;Accession: A69643
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-449 <KUN>
A;Experimental source: strain 168
A;Cross-references: GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13147.1; PID:g2633644
C;Genetics:
A;Gene: htrA
C;Superfamily: proteinase hhoB

alignment_scores:
        Quality: 11.00   Length: 11
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x A69643

Align seg 1/1 to: A69643 from: 1 to: 449   ..

    718 GCCATCAATCCGGGCAATTCCGGCGGCCCGCTG 750
        |||||||||||||||||||||||||||||||||
    284 AlaIleAsnProGlyAsnSerGlyGlyProLeu 294

seq_name: pir2:S77538

seq_documentation_block:
serine proteinase (EC 3.4.21.-) htrA - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein slr1204
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
s.
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
DNA Res. 3, 109-136, 1996
A;Reference number: S74322; MUID:97061201
A;Accession: S77538
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-452 <KAN>
A;Cross-references: EMBL:D90905; GB:AB001339; NID:g1652360; GB:AB001339; PIDN:BAA17385.1; PID:g1652246
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: htrA
C;Superfamily: proteinase hhoB
C;Keywords: hydrolase; serine proteinase

alignment_scores:
        Quality: 11.00   Length: 11
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x S77538

Align seg 1/1 to: S77538 from: 1 to: 452   ..

    718 GCCATCAATCCGGGCAATTCCGGCGGCCCGCTG 750
        |||||||||||||||||||||||||||||||||

---

290 AlaIleAsnProGlyAsnSerGlyGlyProLeu 300

seq_name: pir1:B70426

seq_documentation_block:
periplasmic serine proteinase (EC 3.4.21.-) - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: B70426
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.,
V.
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:9819666
A;Accession: B70426
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-453 <AQF>
A;Cross-references: GB:AE000741; GB:AE000657; NID:g2983841; PIDN:AAC07399.1; PID:g298
A;Experimental source: strain VF5
C;Genetics:
A;Gene: htrA
C;Superfamily: Helicobacter serine proteinase
C;Keywords: hydrolase; serine proteinase

alignment_scores:
        Quality: 11.00   Length: 11
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x B70426

Align seg 1/1 to: B70426 from: 1 to: 453   ..

    718 GCCATCAATCCGGGCAATTCCGGCGGCCCGCTG 750
        |||||||||||||||||||||||||||||||||
    199 AlaIleAsnProGlyAsnSerGlyGlyProLeu 209

seq_name: pir2:F72359

seq_documentation_block:
periplasmic serine proteinase Do (EC 3.4.21.-) - Thermotoga maritima (strain MSB8)
N;Alternate names: heat shock protein htrA
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72359
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316
A;Accession: F72359
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459 <ARN>
A;Cross-references: GB:AE001732; GB:AE000512; NID:g4981087; PIDN:AAD35656.1; PID:g498
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0571
C;Superfamily: Helicobacter serine proteinase
C;Keywords: hydrolase; serine proteinase

alignment_scores:
        Quality: 11.00   Length: 11
        Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x F72359

Align seg 1/1 to: F72359 from: 1 to: 459

```
718 GCCATCAATCCGGGCAATTCCGGCGGCGCCCGCTG 750
    |||||||||||||||||||||||||||||||||||
200 AlaIleAsnProGlyAsnSerGlyGlyProLeu 210
```

seq_name: pir2:G81528

seq_documentation_block:
serine proteinase, HtrA/DegQ/DegS family CP0877 (imported) - Chlamydophila pneumoniae (s
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 21-Jul-2000
C;Accession: G81528
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255
A;Accession: G81528
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-488 <REA>
A;Cross-references: GB:AE002246; GB:AE002161; NID:g7189785; PIDN:AAF38665.1; PID:g718978
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0877
C;Superfamily: Helicobacter serine proteinase

alignment_scores:
    Quality: 11.00   Length: 11
    Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x G81528   ..

Align seg 1/1 to: G81528 from: 1 to: 488

```
718 GCCATCAATCCGGGCAATTCCGGCGGCGCCCGCTG 750
    |||||||||||||||||||||||||||||||||||
232 AlaIleAsnProGlyAsnSerGlyGlyProLeu 242
```

seq_name: pir2:G72011

seq_documentation_block:
do serine proteinase - Chlamydophila pneumoniae (strain CWL029)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C;Accession: G72011
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: G72011
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-488 <ARN>
A;Cross-references: GB:AE001678; GB:AE001363; NID:g4377301; PIDN:AAD19116.1; PID:g437730
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: htrA
C;Superfamily: Helicobacter serine proteinase

alignment_scores:
    Quality: 11.00   Length: 11
    Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x G72011   ..

Align seg 1/1 to: G72011 from: 1 to: 488

```
718 GCCATCAATCCGGGCAATTCCGGCGGCGCCCGCTG 750
    |||||||||||||||||||||||||||||||||||
232 AlaIleAsnProGlyAsnSerGlyGlyProLeu 242
```

seq_name: pir2:B81728

seq_documentation_block:
serine proteinase, HtrA/DegQ/DegS family TC0210 (imported) - Chlamydia muridarum (str
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 21-Jul-2000
C;Accession: B81728
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255
A;Accession: B81728
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-497 <TET>
A;Cross-references: GB:AE002288; GB:AE002160; NID:g7190247; PIDN:AAF39082.1; PID:g719
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0210
C;Superfamily: Helicobacter serine proteinase

alignment_scores:
    Quality: 11.00   Length: 11
    Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x B81728   ..

Align seg 1/1 to: B81728 from: 1 to: 497

```
718 GCCATCAATCCGGGCAATTCCGGCGGCGCCCGCTG 750
    |||||||||||||||||||||||||||||||||||
241 AlaIleAsnProGlyAsnSerGlyGlyProLeu 251
```

seq_name: pir2:H71465

seq_documentation_block:
probable do serine proteinase - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C;Accession: H71465
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A;Reference number: A71570; MUID:99000809
A;Accession: H71465
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-497 <ARN>
A;Cross-references: GB:AE001355; GB:AE001273; NID:g3329292; PIDN:AAC68420.1; PID:g332
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: htrA
C;Superfamily: Helicobacter serine proteinase

alignment_scores:
    Quality: 11.00   Length: 11
    Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x H71465   ..

Fri Mar 23 07:26:30 2001

us-09-388-090-3.oligo7.rpr

Align seg 1/1 to: H71465 from: 1 to: 497

```
718 GCCATCAATCCGGGCAATTCCGGCGGCCCGCTG 750
    |||||||||||||||||||||||||||||||||
241 AlaIleAsnProGlyAsnSerGlyGlyProLeu 251
```

```
OM of: US-09-388-090-3 to: SwissProt_39:*     out_format : pfs

Date: Mar 22, 2001 12:09 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+n2p.model -DEV-xlp
 -o=/cgn2_1/USPTO_spool/US09388090/runat_20032001_092951_25871/app_query.fasta_1.1474
 -DB=SwissProt_39 -QFMT=fastan -SUFFIX=oligo7.rsp -GAPOP=4.500
 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0 -LOOPEXT=0.000
 -QCAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
 -TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR_SCORE=quality
 -THR_MIN=7 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=200000000 -USER=US09388090_@CGN1_1_34 -NCPU=6 -ICPU=3
 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
 Query: US-09-388-090-3
 Query length: 1395
 Database: SwissProt_39:*
 Database sequences: 88757
 Database length: 3294092
 Search time (sec): 26.910000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000
```

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
SwissProt_39:DEGP_ARATH	+	11.00	172.19	0.0130	437	! Q22609 arabidopsis thaliana
SwissProt_39:DEGP_CHLPN	+	11.00	171.48	0.0127	488	! Q9z6t0 chlamydia pneumoniae
SwissProt_39:DEGP_CHLTR	+	11.00	171.36	0.0127	497	! P18584 chlamydia trachomatis hy
SwissProt_39:YYXA_BACSU	+	9.00	136.16	1.44	400	! P39668 bacillus subtilis. hy
SwissProt_39:YHI01_HAEIN	+	8.00	120.98	16.34	247	! P44294 haemophilus influenza
SwissProt_39:TF2B_PYRWO	+	8.00	120.62	16.19	261	! P29095 pyrococcus woesei. tr
SwissProt_39:TF2B_PYRFU	+	8.00	119.72	15.81	300	! Q51731 pyrococcus furiosus.
SwissProt_39:TF2B_PYRHO	+	8.00	119.72	15.81	300	! O59151 pyrococcus horikoshii
SwissProt_39:CYCR_RHOGE	+	8.00	118.43	15.28	366	! P51758 rhodocyclus gelatinos
SwissProt_39:G154_MOUSE	-	8.00	117.41	14.87	429	! P50636 mus musculus (mouse).
SwissProt_39:YIE0_HAEIN	-	8.00	116.91	14.68	463	! P44903 haemophilus influenza
SwissProt_39:SYE_BACST	-	8.00	116.56	14.54	489	! P22249 bacillus stearothermo
SwissProt_39:DEGP_BARHE	+	8.00	116.38	14.47	553	! P54925 bartonella henselae (
SwissProt_39:FXC1_HUMAN	+	8.00	115.76	14.24	553	! Q12948 homo sapiens (human).
SwissProt_39:FXC1_MOUSE	-	8.00	115.76	14.24	553	! Q61572 mus musculus (mouse).
SwissProt_39:Y4BJ_RHISN	+	8.00	114.49	13.77	630	! P55397 rhizobium sp. (strain
SwissProt_39:TF2B_METJA	+	8.00	114.49	13.77	673	! Q58192 methanococcus jannas
SwissProt_39:BGH3_PIG	-	8.00	114.40	13.74	683	! O11780 sus scrofa (pig). trans
SwissProt_39:PQQF_PSEFL	-	8.00	113.15	13.29	829	! P55174 pseudomonas fluoresce
SwissProt_39:AGRI_DISOM	-	8.00	110.10	12.27	1328	! Q90404 discopyge ommata (el
SwissProt_39:PHI3_MYTCA	+	7.00	113.69	228.43	45	! P11860 mytilus californianus
SwissProt_39:CY2_AGRTC	-	7.00	107.85	195.54	111	! P00081 agrobacterium tumefaci
SwissProt_39:YPMC_ECOLI	+	7.00	107.79	195.33	112	! P07115 escherichia coli. hyp
SwissProt_39:EXBD_HAEDU	-	7.00	106.88	190.88	129	! P51809 haemophilus ducreyi.
SwissProt_39:EXBD_PASHA	+	7.00	106.12	187.11	145	! P72203 pasteurella haemolyti
SwissProt_39:YJGX_ECOLI	-	7.00	105.99	186.46	148	! P39349 escherichia coli. hyp
SwissProt_39:VL1_HPV64	-	7.00	105.77	185.41	153	! P50824 human papillomavirus t
SwissProt_39:YSAA_ECOLI	+	7.00	105.61	184.60	157	! P56256 escherichia coli. put
SwissProt_39:YCRS_MXXA	-	7.00	105.09	182.11	170	! Q06929 myxococcus xanthus. h
SwissProt_39:VNSC_CDVO	-	7.00	104.94	181.39	174	! P06941 canine distemper virus
SwissProt_39:CD3E_CHICK	-	7.00	104.37	178.69	175	! Q98910 gallus gallus (chicke
SwissProt_39:COOF_RHORU	+	7.00	104.11	177.44	190	! P31894 rhodospirillum rubrum
SwissProt_39:SDC4_MOUSE	-	7.00	104.07	177.28	198	! O35988 mus musculus (mouse).
SwissProt_39:YE74_HAEIN	+	7.00	103.98	176.83	199	! O57213 haemophilus influenza
SwissProt_39:HIL_MYTTR	+	7.00	103.98	176.83	202	! Q05831 mytilus trossulus (blu
SwissProt_39:SDC4_RAT	-	7.00	103.98	176.83	202	! P34901 rattus norvegicus (rat)
SwissProt_39:Y829_SYNY3	-	7.00	103.66	175.38	212	! Q55423 synechocystis sp. (str
SwissProt_39:NRFC_ECOLI	+	7.00	103.34	173.87	223	! P32708 escherichia coli. nrf
SwissProt_39:Y962_MYCTU	-	7.00	103.31	173.74	224	! P71548 mycobacterium tubercu
SwissProt_39:NRFC_HAEIN	+	7.00	103.28	173.61	225	! P45015 haemophilus influ
SwissProt_39:VATE_DROME	-	7.00	103.25	173.48	226	! P54611 drosophila melano
SwissProt_39:H1_WHEAT	+	7.00	102.97	172.20	236	! P27806 triticum aestivum (
SwissProt_39:DHSB_PORPU	+	7.00	102.89	171.83	239	! P80477 porphyra purpurea
SwissProt_39:YGAZ_ECOLI	-	7.00	102.73	171.11	245	! P77630 escherichia coli.
SwissProt_39:OCCP_AGRT6	-	7.00	102.29	169.16	262	! P35117 agrobacterium tume
SwissProt_39:ERF1_ARATH	-	7.00	102.20	168.72	266	! O80337 arabidopsis thali
SwissProt_39:Y890_AQUAE	-	7.00	102.17	168.62	267	! O67039 aquifex aeolicus
SwissProt_39:TRPA_MYCIT	-	7.00	102.08	168.19	271	! O68906 mycobacterium int
SwissProt_39:LEP4_BURPS	+	7.00	101.98	167.77	275	! Q9zf70 burkholderia pseu
SwissProt_39:YARA_PROST	+	7.00	101.82	167.05	282	! P46117 providencia stuari
SwissProt_39:YG53_YEAST	-	7.00	101.68	166.45	288	! P50091 saccharomyces cere
SwissProt_39:PFLC_ECOLI	+	7.00	101.59	166.06	292	! P32675 escherichia coli.
SwissProt_39:MOXJ_METEX	+	7.00	101.42	165.30	300	! P16028 methylobacterium (
SwissProt_39:DDL_BUCAP	-	7.00	101.29	164.74	306	! O51927 buchnera aphidicol
SwissProt_39:Y205_AQUAE	-	7.00	101.29	164.74	306	! O66401 aquifex aeolicus
SwissProt_39:MOCB_RHIME	-	7.00	101.23	164.47	309	! P49308 rhizobium meliloti
SwissProt_39:YHC1_BACSU	-	7.00	101.14	164.11	313	! P54593 bacillus subtilis.
SwissProt_39:PPAC_DEIRA	-	7.00	101.12	164.02	314	! Q9rrb7 deinococcus radio
SwissProt_39:NECD_MOUSE	+	7.00	100.90	163.06	325	! P25233 mus musculus (mous
SwissProt_39:ADH1_BACST	-	7.00	100.67	162.05	337	! P12311 bacillus stearothe
SwissProt_39:G3PX_HORVU	-	7.00	100.67	162.05	337	! P26517 hordeum vulgare (t
SwissProt_39:ADH3_BACST	+	7.00	100.63	161.89	339	! P43328 bacillus stearothe
SwissProt_39:YI21_SYNY3	+	7.00	100.13	159.79	366	! P73714 synechocystis sp.
SwissProt_39:ALF_PLAFA	-	7.00	100.08	159.57	369	! P14223 plasmodium falcipa
SwissProt_39:P2A1_YEAST	-	7.00	100.08	159.57	369	! P23594 saccharomyces cere
SwissProt_39:TMAF_AVIS4	-	7.00	100.08	159.57	369	! P23091 avian myeloblastos
SwissProt_39:T186_ECOLI	-	7.00	100.06	159.49	370	! P08409 escherichia coli.
SwissProt_39:T421_ECOLI	+	7.00	100.04	159.42	371	! P11901 escherichia coli.
SwissProt_39:YO28_ARCFU	-	7.00	100.04	159.42	371	! O30207 archaeoglobus fulg
SwissProt_39:P2A2_YEAST	-	7.00	99.94	158.98	377	! P23595 saccharomyces cere
SwissProt_39:ENTC_ECOLI	+	7.00	99.71	158.00	391	! P10377 escherichia coli.
SwissProt_39:ODBB_BOVIN	-	7.00	99.69	157.93	392	! P21839 bos taurus (bovine
SwissProt_39:DOM3_CAEEL	-	7.00	99.67	157.86	393	! Q10660 caenorhabditis ele
SwissProt_39:PHHY_PSEAE	+	7.00	99.66	157.79	394	! P20586 pseudomonas aerugi
SwissProt_39:PHHY_PSEFL	+	7.00	99.66	157.79	394	! P00438 pseudomonas fluoro
SwissProt_39:PAR2_HUMAN	+	7.00	99.61	157.59	397	! P55085 homo sapiens (huma
SwissProt_39:PAR2_RAT	+	7.00	99.61	157.59	397	! Q63645 rattus norvegicus (r
SwissProt_39:PAR2_MOUSE	+	7.00	99.57	157.45	399	! P55086 mus musculus (mous
SwissProt_39:COAT_BBV	+	7.00	99.45	156.92	407	! P04329 black beetle virus (
SwissProt_39:YPEB_SYNP2	+	7.00	99.38	156.66	411	! P42784 synechococcus sp.
SwissProt_39:VGLD_HSVBP	-	7.00	99.29	156.27	417	! P24906 bovine herpesvirus
SwissProt_39:VGLD_HSVBS	-	7.00	99.29	156.27	417	! Q08100 bovine herpesvirus
SwissProt_39:Y824_DEIRA	-	7.00	99.17	155.77	418	! Q9rw45 deinococcus radiod
SwissProt_39:IDH_HELPY	+	7.00	99.17	155.77	425	! P56063 helicobacter pylori
SwissProt_39:RP54_RHOCA	+	7.00	99.15	155.71	426	! P0P433 rhodobacter capsul
SwissProt_39:YAEL_ECOLI	+	7.00	98.80	154.26	450	! P37764 escherichia coli.
SwissProt_39:DEGQ_ECOLI	+	7.00	98.71	153.97	455	! P39099 escherichia coli.
SwissProt_39:RFK9_ECOLI	-	7.00	98.71	153.91	456	! P37755 escherichia coli.
SwissProt_39:CYSG_SALTY	-	7.00	98.57	153.85	457	! P25924 s siroheme synthas
SwissProt_39:HEMN_SYNY3	-	7.00	98.49	153.34	466	! P74132 synechocystis sp.
SwissProt_39:YUB7_MYCTU	-	7.00	98.30	153.01	472	! O53304 mycobacterium tube
SwissProt_39:BAF1_KLULA	+	7.00	98.18	152.25	486	! P26375 kluyveromyces lact
SwissProt_39:CRTD_RHOSH	+	7.00	98.11	151.77	495	! Q01671 rhodobacter sphaer
SwissProt_39:AMPA_BACSU	-	7.00	97.97	151.51	500	! O32106 bacillus subtilis.
SwissProt_39:PGKA_CRIFA	+	7.00	97.95	151.26	505	! P25055 crithidia fascicul
SwissProt_39:GUN0_PSEFL	+	7.00	97.76	150.95	511	! P18126 pseudomonas fluore
SwissProt_39:DEGP_RICPR	+	7.00	97.75	150.85	513	! O05942 rickettsia prowaze
SwissProt_39:VL1_HPV34	+	7.00	97.76	150.11	528	! P36738 human papillomaviru
SwissProt_39:YJCC_ECOLI	-	7.00	97.75	150.06	528	! P32701 escherichia coli.
SwissProt_39:SP15_STRGR	-	7.00	97.74	150.01	529	! P19471 streptomyces grise
SwissProt_39:YA9A_SCHPO	+	7.00	97.68	149.82	530	! Q09788 schizosaccharomyce
SwissProt_39:ML51_HUMAN	-	7.00	97.68	149.77	534	! O15234 homo sapiens (huma
SwissProt_39:YMA2_CAEEL	+	7.00	97.67	149.73	535	! P34447 caenorhabditis ele
SwissProt_39:VNCA_AAV2	+	7.00	97.65	149.68	536	! P03132 adeno-associated vi
SwissProt_39:CPF6_RAT	-	7.00	97.63	149.58	537	! P51871 rattus norvegicus (
SwissProt_39:COAT_LORDV	+	7.00	97.63	149.44	539	! P54635 lordsdale virus (h
SwissProt_39:YM87_MYCTU	+	7.00	97.59	149.30	542	! Q50678 mycobacterium tube
SwissProt_39:ATPA_YEAST	-	7.00	97.56	149.26	545	! P07251 saccharomyces cere
SwissProt_39:PUT2_AGABI	+	7.00	97.55	149.26	546	! P78568 agaricus bisporus
SwissProt_39:DEAD_MYCTU	-	7.00	97.35	148.48	563	! Q11039 mycobacterium tube
SwissProt_39:GUNB_CLOTM	+	7.00	97.35	148.48	563	! P04956 clostridium thermo
SwissProt_39:AMP2_LYCES	-	7.00	97.28	148.21	569	! Q42876 lycopersicon escul

SwissProt_39:URE1_SYNY3	+	7.00	97.28	148.21	569	P73061 synechocystis sp.(st
SwissProt_39:AMP1_LYCES	-	7.00	97.26	148.12	571	Q10712 lycopersicon esculent
SwissProt_39:AMP1_SOLTU	-	7.00	97.23	148.03	573	P31427 solanum tuberosum (po
SwissProt_39:FHAC_BORPE	+	7.00	97.11	147.55	584	P35077 bordetella pertussis
SwissProt_39:EYTA_CAEEL	+	7.00	97.04	147.30	590	P34322 caenorhabditis elegan
SwissProt_39:EYAL_MOUSE	-	7.00	97.03	147.25	591	P97767 mus musculus (mouse)
SwissProt_39:YG2S_YEAST	-	7.00	96.88	146.67	605	P53261 saccharomyces cerevis
SwissProt_39:YG3F_YEAST	-	7.00	96.79	146.30	614	P53283 saccharomyces cerevis
SwissProt_39:ILVD_STRCO	-	7.00	96.75	146.18	617	O69198 streptomyces coelicol
SwissProt_39:CA28_HUMAN	+	7.00	96.57	145.46	635	P25067 homo sapiens (human)
SwissProt_39:DNAK_FRATU	+	7.00	96.50	145.19	642	P48205 francisella tularensi
SwissProt_39:NEU1_RABIT	-	7.00	96.45	145.00	647	P40489 oryctolagus cuniculus
SwissProt_39:YDEE_SCHPO	+	7.00	96.44	144.96	648	Q10447 schizosaccharomyces
SwissProt_39:MI0_CAEEL	+	7.00	96.42	144.88	650	P34400 caenorhabditis elegan
SwissProt_39:GNP1_YEAST	-	7.00	96.29	144.40	663	P48813 saccharomyces cerevis
SwissProt_39:FLAY_CAUCR	+	7.00	96.14	143.81	679	P15346 caulobacter crescentu
SwissProt_39:NEU1_RABIT	-	7.00	95.90	142.93	704	P42675 oryctolagus cuniculus
SwissProt_39:YCCS_ECOLI	+	7.00	95.76	142.38	720	P75870 escherichia coli. hyp
SwissProt_39:SUF_DROME	+	7.00	95.64	141.95	733	P25991 drosophila melanogaste
SwissProt_39:FCUA_YEREN	+	7.00	95.42	141.14	758	Q9Z2X6 yersinia enterocoliti
SwissProt_39:MUTS_THECA	-	7.00	94.94	139.34	817	Q9ZIX6 thermus aquaticus (su
SwissProt_39:ATCU_RHILV	+	7.00	94.75	138.66	841	Q95V3 rhizobium leguminosar
SwissProt_39:VGLH_HSVBC	-	7.00	94.74	138.63	842	P27599 bovine herpesvirus ty
SwissProt_39:DEXT_STRMU	-	7.00	94.68	138.41	850	Q54443 streptococcus mutans.
SwissProt_39:SSY1_YEAST	-	7.00	94.67	138.35	852	Q03770 saccharomyces cerevis
SwissProt_39:OSH1_YEAST	-	7.00	94.61	138.16	859	P35845 saccharomyces cerevis
SEQUENCE OF FPRP RAT +		7.00	94.47	137.62	879	Q62786 rattus norvegicus (rat)
SwissProt_39:ITH3_MOUSE	+	7.00	94.41	137.43	886	Q61704 mus musculus (mouse)
SwissProt_39:DPO1_BORBU	-	7.00	94.26	136.86	908	Q51498 borrelia burgdorferi
SwissProt_39:ST20_YEAST	-	7.00	94.04	136.08	939	Q03087 saccharomyces cerevis
SwissProt_39:AMYG_CANAL	+	7.00	93.99	135.90	946	O74254 candida albicans (yea
SwissProt_39:KF5C_MOUSE	+	7.00	93.92	135.66	956	Q60738 mus musculus (mouse).
SwissProt_39:PTPN_MOUSE	+	7.00	93.77	135.11	979	Q60673 mus musculus (mouse).
SwissProt_39:PTPN_RAT	+	7.00	93.74	135.02	983	Q63259 rattus norvegicus (rat)
SwissProt_39:YPT7_CAEEL	+	7.00	93.50	134.15	1021	P41885 caenorhabditis elega
SwissProt_39:ANGR_VIBAN	+	7.00	93.55	133.55	1048	P19828 vibrio anguillarum.
SwissProt_39:ITA5_HUMAN	+	7.00	93.32	133.53	1049	P08648 homo sapiens (human)
SwissProt_39:ITA5_XENLA	+	7.00	93.32	133.51	1050	Q06274 xenopus laevis (afri
SwissProt_39:ITA5_MOUSE	+	7.00	93.30	133.44	1053	P11688 mus musculus (mouse)
SwissProt_39:PLD1_HUMAN	+	7.00	93.17	133.00	1074	Q13393 homo sapiens (human)
SwissProt_39:PLD1_RAT	-	7.00	93.16	132.97	1075	P70496 rattus norvegicus (rat
SwissProt_39:TRA1_CAEEL	+	7.00	92.96	132.25	1110	P34708 caenorhabditis elega
SwissProt_39:YN96_YEAST	-	7.00	92.92	132.11	1117	P53753 saccharomyces cerevi
SwissProt_39:MAP4_MOUSE	+	7.00	92.87	131.95	1125	P27546 mus musculus (mouse)
SwissProt_39:MAP4_HUMAN	+	7.00	92.72	131.41	1152	P27816 homo sapiens (human)
SwissProt_39:NKFJ_ENTAG	+	7.00	92.60	131.01	1173	P19543 enterobacter agglome
SwissProt_39:NKC1_SQUAC	+	7.00	92.50	130.67	1191	P55013 squalus acanthias (s
SwissProt_39:RPOB_THEAC	+	7.00	92.48	130.60	1195	Q05587 thermoplasma acidoph
SwissProt_39:YMZ9_YEAST	-	7.00	92.35	130.15	1219	Q04439 saccharomyces cerevi
SwissProt_39:DNBI_HCMVA	+	7.00	92.27	129.78	1235	P17147 human cytomegaloviru
SwissProt_39:TOP2_LEICH	+	7.00	92.24	129.74	1240	Q01078 leishmania chagasi.
SwissProt_39:TBP7_CAEEL	+	7.00	92.23	129.74	1242	P54816 caenorhabditis elega
SwissProt_39:MYS3_YEAST	-	7.00	92.07	129.20	1273	P36006 saccharomyces cerevi
SwissProt_39:VCAP_HSV11	+	7.00	91.58	127.52	1374	P06491 herpes simplex virus
SwissProt_39:MHP1_YEAST	-	7.00	91.46	127.15	1398	P43638 saccharomyces cerevi
SwissProt_39:YKB4_YEAST	-	7.00	91.35	126.78	1422	Q98936 gallus gallus (chick
SwissProt_39:PTPG_MOUSE	+	7.00	91.26	126.48	1442	Q05909 mus musculus (mouse)
SwissProt_39:PTPG_HUMAN	+	7.00	91.25	126.43	1445	P23470 homo sapiens (human)
SwissProt_39:KDGE_DROME	+	7.00	91.21	126.30	1454	Q09103 drosophila melanogas
SwissProt_39:N153_RAT	+	7.00	91.15	126.09	1468	P49791 rattus norvegicus (rat
SwissProt_39:PTP1_DROME	+	7.00	90.47	123.86	1630	P35992 drosophila melanogas
SwissProt_39:CA64_YEAST	-	7.00	90.28	123.25	1678	Q14031 homo sapiens (human)
SwissProt_39:CYAA_SCHPO	+	7.00	90.23	123.08	1692	P14605 schizosaccharomyces
SwissProt_39:CCAC_MOUSE	+	7.00	89.96	122.21	1764	P34241 saccharomyces cerevi
SwissProt_39:CCAC_RAT	+	7.00	88.71	118.26	2139	Q01815 mus musculus (mouse)
SEQUENCE OF CCAC_RAT +		7.00	88.62	117.97	2169	P22002 rattus norvegicus (rat
SwissProt_39:CCAC_RABIT	+	7.00	88.62	117.96	2171	P15381 oryctolagus cuniculu
SwissProt_39:STCA_EMENI	+	7.00	88.59	117.86	2181	Q12397 emericella nidulans
SwissProt_39:CCAC_HUMAN	+	7.00	88.47	117.50	2221	Q13936 homo sapiens (human)
SwissProt_39:2AN_PIG	-	7.00	87.77	115.34	2476	Q28983 sus scrofa (pig). zonad
SwissProt_39:CCAA_HUMAN	+	7.00	87.69	115.11	2505	P00555 homo sapiens (human)
SwissProt_39:KI67_HUMAN	+	7.00	86.00	110.08	3256	P46013 homo sapiens (human)
SwissProt_39:POLG_MCFA	+	7.00	85.83	109.60	3341	P33515 m genome polyprotein

SwissProt_39:OL56_STRAT	-	7.00	85.49	108.63	3519	Q07017 streptomyces antl
SwissProt_39:LMI2_CAEEL	-	7.00	85.22	107.85	3672	Q21313 caenorhabditis el
SwissProt_39:LRP1_HUMAN	-	7.00	83.84	104.00	4544	Q07954 homo sapiens (hum

seq_name: SwissProt_39:DEGP_ARATH

seq_documentation_block:
ID   DEGP_ARATH   STANDARD;   PRT;   437 AA.
AC   O22630;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROTEASE DO-LIKE PRECURSOR (EC 3.4.21.-).
GN   DEGP.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX   MEDLINE=98175982; PubMed=9507020;
RA   Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
RT   "Identification and characterization of DegP, a serine protease
RT   associated with the luminal side of the thylakoid membrane.";
RL   J. Biol. Chem. 273:7094-7098(1998).
RN   [2]
RP   SEQUENCE OF 104-118.
RA   Kieselbach T., Bystedt M., Schroeder W.P.;
RL   Submitted (JUL-2000) to the SWISS-PROT data bank.
CC   -!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
CC   MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
CC   CAN DEGRADE BETA-CASEIN.
CC   -!- ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
CC   O-PHENANTHROLINE.
CC   -!- SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
CC   MEMBRANE.
CC   -!- INDUCTION: HEAT SHOCK.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C; ALSO KNOWN AS THE
CC   DEGP/DEGQ/DEGS FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF028842; AAC39436.1; -.
DR   MEROPS; S01.279; -.
DR   MENDEL; 25267; Arath;3069;25267.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 1.
DR   PRINTS; PR00834; PROTEASES2C.
KW   Hydrolase; Serine protease; Transit peptide; Chloroplast.
FT   TRANSIT   1   103   CHLOROPLAST.
FT   CHAIN   104   437   PROTEASE DO-LIKE.
FT   ACT_SITE   171   171   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   201   201   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   280   280   CHARGE RELAY SYSTEM (POTENTIAL).
SQ   SEQUENCE   437 AA; 46213 MW; 1497B1AB3F5FF2A4 CRC64;

alignment_scores:
     Quality: 11.00   Length: 11
     Ratio: 1.000   Gaps: 0
     Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x DEGP_ARATH

seq_name: SwissProt_39:DEGP_CHLTR

seq_documentation_block:
ID   DEGP_CHLTR   STANDARD;   PRT;   497 AA.
AC   P18584; O84830;
DT   01-NOV-1990 (Rel. 16, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROBABLE SERINE PROTEASE DO-LIKE PRECURSOR (EC 3.4.21.-) (59 KDA
DE   IMMUNOGENIC PROTEIN) (SK59).
GN   DEGP OR HTRA OR CT823.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SEROVAR L2;
RX   MEDLINE=90337348; PubMed=2379836;
RA   Kahane S., Weinstein Y., Sarov I.;
RT   "Cloning, characterization and sequence of a novel 59-kDa protein of
RT   Chlamydia trachomatis.";
RL   Gene 90:61-67(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/CX;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C; ALSO KNOWN AS THE
CC   DEGP/DEGQ/DEGS FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.
CC   -!- CAUTION: REF.1 SEQUENCE WAS INCORRECT: DUE TO SEQUENCING ERRORS
CC   THE AUTHORS TRANSLATED THEIR PUTATIVE 59 KDA IMMUNOGENIC PROTEIN
CC   ON THE COMPLEMENTARY STRAND TO THAT OF WHAT SEEMS TO BE THE REAL
CC   PROTEIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AE001355; AAC68420.1; -.
DR   EMBL; M31119; AAA23116.1; -.
DR   INTERPRO; IPR000126; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00595; PDZ; 2.
DR   PFAM; PF00089; trypsin; 1.
DR   PRINTS; PR00834; PROTEASES2C.
DR   PRINTS; PR00839; V8PROTEASE.
KW   Hydrolase; Serine protease; Signal; Antigen.
FT   SIGNAL   1   16   POTENTIAL.
FT   CHAIN   17   497   PROBABLE SERINE PROTEASE DO-LIKE.
FT   DOMAIN   128   289   CATALYTIC.
FT   DOMAIN   290   381   PDZ.
FT   DOMAIN   394   485   PDZ.
FT   ACT_SITE   143   143   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   173   173   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   247   247   CHARGE RELAY SYSTEM (POTENTIAL).
SQ   SEQUENCE   497 AA; 53244 MW; 86A5E31BB84A38BA CRC64;

alignment_scores:
        Quality: 11.00   Length: 11
          Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

Align seq 1/1 to: DEGP_ARATH from: 1 to: 437

718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
    ||||||||||||||||||||||||||||||||
274 AlaIleAsnProGlyAsnSerGlyGlyProLeu 284

seq_name: SwissProt_39:DEGP_CHLPN

seq_documentation_block:
ID   DEGP_CHLPN   STANDARD;   PRT;   488 AA.
AC   Q9Z6T0;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROBABLE SERINE PROTEASE DO-LIKE PRECURSOR (EC 3.4.21.-).
GN   DEGP OR HTRA OR CPN0979
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C; ALSO KNOWN AS THE
CC   DEGP/DEGQ/DEGS FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AE001678; AAD19116.1; -.
DR   MEROPS; S01.273; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 2.
DR   PRINTS; PR00834; PROTEASES2C.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL   1   20   POTENTIAL.
FT   CHAIN   21   488   PROBABLE SERINE PROTEASE DO-LIKE.
FT   DOMAIN   119   280   CATALYTIC.
FT   DOMAIN   281   372   PDZ.
FT   DOMAIN   388   476   PDZ.
FT   ACT_SITE   134   134   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   164   164   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   238   238   CHARGE RELAY SYSTEM (POTENTIAL).
SQ   SEQUENCE   488 AA; 52311 MW; 0EE7E0F88F106F49 CRC64;

alignment_scores:
        Quality: 11.00   Length: 11
          Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x DEGP_CHLPN   ..

Align seq 1/1 to: DEGP_CHLPN from: 1 to: 488

718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
    ||||||||||||||||||||||||||||||||
232 AlaIleAsnProGlyAsnSerGlyGlyProLeu 242

```
alignment_block:
US-09-388-090-3 x DEGP_CHLTR

Align seg 1/1  to: DEGP_CHLTR  from: 1  to: 497

      718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
          |||||||||||||||||||||||||||||||||||
      241 AlaIleAsnProGlyAsnSerGlyGlyProLeu 251

seq_name: SwissProt_39:YYXA_BACSU

seq_documentation_block:
ID  YYXA_BACSU   STANDARD;    PRT;   400 AA.
AC  P39668;
DT  01-FEB-1995 (Rel. 31, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  HYPOTHETICAL PROTEASE IN ROCR-PURA INTERGENIC REGION (EC 3.4.21.-).
GN  YYXA OR YYCK.
OS  Bacillus subtilis.
OC  Bacteria; Firmicutes; Bacillus/Clostridium group;
OC  Bacillus/Staphylococcus group; Bacillus.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-168.
RA  Kasahara Y., Nakai S., Yoshikawa H., Ogasawara N.;
RL  Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN  [2]
RP  SEQUENCE OF 317-400 FROM N.A.
RC  STRAIN-168 / MARBURG;
RX  MEDLINE=94156824; PubMed=8113162;
RA  Calogero S., Gardan R., Glaser P., Schweitzer J., Rapoport G.,
RA  Debarbouille M.;
RT  "RocR, a novel regulatory protein controlling arginine utilization in
RT  Bacillus subtilis, belongs to the NtrC/NifA family of transcriptional
RT  activators.";
RL  J. Bacteriol. 176:1234-1241(1994).
CC  -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C; ALSO KNOWN AS THE
CC      DEGP/DEGQ/DEGS FAMILY.
CC  -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
CC  -------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -------------------------------------------------------------------------
DR  EMBL; D78193; BAA11295.1; -.
DR  EMBL; L22006; -; NOT_ANNOTATED_CDS.
DR  EMBL; Z99124; CAB16073.1; -.
DR  MEROPS; S01.273; -.
DR  SUBTILIST; BG11054; YYXA.
DR  INTERPRO; IPR001254; -.
DR  INTERPRO; IPR001478; -.
DR  INTERPRO; IPR001940; -.
DR  PFAM; PF00595; PDZ; 1.
DR  PFAM; PF00089; trypsin; 1.
DR  PRINTS; PR00834; PROTEASES2C.
KW  Hypothetical protein; Hydrolase; Serine protease.
FT  ACT_SITE  129  129  CHARGE RELAY SYSTEM (POTENTIAL).
FT  ACT_SITE  159  159  CHARGE RELAY SYSTEM (POTENTIAL).
FT  ACT_SITE  243  243  CHARGE RELAY SYSTEM (POTENTIAL).
SQ  SEQUENCE  400 AA;  42788 MW;  91A6E0E897B6F64C CRC64;

alignment_scores:
        Quality:  9.00          Length:  9
          Ratio:  1.000            Gaps:  0
Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
US-09-388-090-3 x YYXA_BACSU

Align seg 1/1  to: YYXA_BACSU  from: 1  to: 400

      718 GCCATCAATCCGGGCAATTCCGGCGCGGC 744
          |||||||||||||||||||||||||||||
      237 AlaIleAsnProGlyAsnSerGlyGly 245

seq_name: SwissProt_39:YH01_HAEIN

seq_documentation_block:
ID  YH01_HAEIN   STANDARD;    PRT;   247 AA.
AC  P44292;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  HYPOTHETICAL PROTEIN HI1701
GN  HI1701.
OS  Haemophilus influenzae.
OC  Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC  Haemophilus.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-RD / KW20;
RX  MEDLINE=95350630; PubMed-7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA  McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA  Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA  Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  "Whole-genome random sequencing and assembly of Haemophilus
RT  influenzae Rd.";
RL  Science 269:496-512(1995).
CC  -------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -------------------------------------------------------------------------
DR  EMBL; U32843; AAC23347.1; -.
DR  TIGR; HI1701; -.
KW  Hypothetical protein; Transmembrane.
FT  TRANSMEM  11  31  POTENTIAL.
FT  TRANSMEM  39  59  POTENTIAL.
SQ  SEQUENCE  247 AA;  27612 MW;  8242407AB6F8D739 CRC64;

alignment_scores:
        Quality:  8.00          Length:  8
          Ratio:  1.000            Gaps:  0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-388-090-3/rev x YH01_HAEIN

Align seg 1/1  to: YH01_HAEIN  from: 1  to: 247

      575 GGATTGCCGATTTGACGACGGGT 552
          |||||||||||||||||
      118 GlyLeuProIleLeuThrThrGly 125

seq_name: SwissProt_39:TF2B_PYRWO

seq_documentation_block:
ID  TF2B_PYRWO   STANDARD;    PRT;   261 AA.
```

```
AC   P29095;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRANSCRIPTION INITIATION FACTOR IIB HOMOLOG (TFIIB).
GN   TFB.
OS   Pyrococcus woesei.
OC   Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93324384; PubMed=8332505;
RA   Creti R., Londei P., Cammarano P.;
RT   "Complete nucleotide sequence of an archaeal (Pyrococcus woesei) gene
RT   encoding a homolog of eukaryotic transcription factor IIB (TFIIB).";
RL   Nucleic Acids Res. 21:2942-2942(1993).
RN   [2]
RP   SEQUENCE OF 110-261 FROM N.A.
RX   MEDLINE=92130258; PubMed=1723106;
RA   Creti R., Citarella F., Tiboni O., Sanangelantoni A.M., Palm P.,
RA   Cammarano P.;
RT   "Nucleotide sequence of a DNA region comprising the gene for
RT   elongation factor 1 alpha (EF-1 alpha) from the ultrathermophilic
RT   archaeote Pyrococcus woesei: phylogenetic implications.";
RL   J. Mol. Evol. 33:332-342(1991).
RN   [3]
RP   SIMILARITY TO EUKARYOTIC TFIIB.
RX   MEDLINE=93046619; PubMed=1423586;
RA   Ouzounis C., Sander C.;
RT   "TFIIB, an evolutionary link between the transcription machineries of
RT   archaebacteria and eukaryotes.";
RL   Cell 71:189-190(1992).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 62-261.
RX   MEDLINE=97322320; PubMed=9177165;
RA   Kosa P.F., Ghosh G., Dedecker B.S., Sigler P.B.;
RT   "The 2.1-A crystal structure of an archaeal preinitiation complex:
RT   TATA-box-binding protein/transcription factor (II)B core/TATA-box.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:6042-6047(1997).
CC   -!- FUNCTION: STABILIZES TBP BINDING TO AN ARCHAEAL BOX-A PROMOTER.
CC       ALSO RESPONSIBLE FOR RECRUITING RNA POLYMERASE II TO THE PRE-
CC       INITIATION COMPLEX (DNA-TBP-TFIIB) (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TFIIB FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; X70668; CAA50006.1; -.
DR   EMBL; X59857; -; NOT_ANNOTATED_CDS.
DR   PIR; S31706; S31706.
DR   PDB; 1AIS; 07-JUL-97.
DR   INTERPRO; IPR000812; -.
DR   PFAM; PF00382; transcript_fac2; 2.
DR   PROSITE; PS00782; TFIIB; 2.
KW   Transcription regulation; Repeat; 3D-structure.
FT   REPEAT       80    156
FT   REPEAT      176    252
SQ   SEQUENCE   261 AA;  29748 MW;  21F1493CDEFADB64 CRC64;

alignment_scores:
        Quality:      8.00            Length:      8
          Ratio:     1.000             Gaps:       0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x TF2B_PYRWO

Align seg 1/1  to: TF2B_PYRWO  from: 1  to: 261
```

```
1318 CGGGTTCCGACGCGGCAGAACGC 1341
     |||||||||||||||||||||||
  61 ArgValSerAspAlaAlaGluArg 68

seq_name: SwissProt_39:TF2B_PYRFU

seq_documentation_block:
ID   TF2B_PYRFU     STANDARD;      PRT;     300 AA.
AC   Q51731;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRANSCRIPTION INITIATION FACTOR IIB HOMOLOG (TFIIB).
GN   TFB.
OS   Pyrococcus furiosus.
OC   Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zeng Q., Lewis L.M., Colangelo C.M., Dong J., Scott R.A.;
RT   "A transcription factor (TFIIB) homolog from the hyperthermophilic
RT   archaeon Pyrococcus furiosus binds Zn or Fe in an N-terminal Cys4
RT   motif.";
RL   J. Biol. Inorg. Chem. 1:162-168(1996).
RN   [2]
RP   STRUCTURE BY NMR OF 1-50.
RX   MEDLINE=9616429; PubMed=8854536;
RA   Zhu W., Zeng Q., Colangelo C.M., Lewis L.M., Summers M.F.,
RA   Scott R.A.;
RT   "The N-terminal domain of TFIIB from Pyrococcus furiosus forms a zinc
RT   ribbon.";
RL   Nat. Struct. Biol. 3:122-124(1996).
CC   -!- FUNCTION: STABILIZES TBP BINDING TO AN ARCHAEAL BOX-A PROMOTER.
CC       ALSO RESPONSIBLE FOR RECRUITING RNA POLYMERASE II TO THE PRE-
CC       INITIATION COMPLEX (DNA-TBP-TFIIB).
CC   -!- SIMILARITY: BELONGS TO THE TFIIB FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; U48391; AAC43724.1; -.
DR   PDB; 1PFT; 17-AUG-96.
DR   INTERPRO; IPR000812; -.
DR   PFAM; PF00382; transcript_fac2; 2.
DR   PRINTS; PR00685; TIFACTORIIB.
DR   PROSITE; PS00782; TFIIB; 2.
KW   Transcription regulation; Repeat; Zinc-finger; 3D-structure.
FT   ZN_FING       7     29
SQ   SEQUENCE   300 AA;  34105 MW;  476D7CA32B2ED4C1 CRC64;

alignment_scores:
        Quality:      8.00            Length:      8
          Ratio:     1.000             Gaps:       0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x TF2B_PYRFU

Align seg 1/1  to: TF2B_PYRFU  from: 1  to: 300

1318 CGGGTTCCGACGCGGCAGAACGC 1341
     |||||||||||||||||||||||
 100 ArgValSerAspAlaAlaGluArg 107

seq_documentation_block:
```

```
ID   TF2B_PYRHO     STANDARD;      PRT;   300 AA.
AC   O59151;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRANSCRIPTION INITIATION FACTOR IIB HOMOLOG (TFIIB).
GN   TFB OR PH1482.
OS   Pyrococcus horikoshii.
OC   Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OT3;
RX   MEDLINE=98344137; PubMed=9679194;
RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA   Masuchi Y., Shizuya H., Kikuchi H.;
RT   "Complete sequence and gene organization of the genome of a hyper-
RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL   DNA Res. 5:55-76(1998).
CC   -!- FUNCTION: STABILIZES TBP BINDING TO AN ARCHAEAL BOX-A PROMOTER.
CC       ALSO RESPONSIBLE FOR RECRUITING RNA POLYMERASE II TO THE PRE-
CC       INITIATION COMPLEX (DNA-TBP-TFIIB) (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TFIIB FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use  by non-profit institutions as long  as its content is  in  no  way
CC   modified and this statement is not removed.  Usage  by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AP000006; BAA30589.1; -.
DR   INTERPRO; IPR000812; -.
DR   PFAM; PF00382; transcript_fac2; 2.
DR   PRINTS; PR00685; TIFACTORIIB.
DR   PROSITE; PS00782; TFIIB; 2.
KW   Transcription regulation; Repeat; Zinc-finger.
FT   ZN_FING       7     29
SQ   SEQUENCE  300 AA;  34097 MW;  DE9758F398BC855F CRC64;
```

alignment_scores:
          Quality: 8.00   Length: 8
          Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x TF2B_PYRHO

Align seg 1/1 to: TF2B_PYRHO from: 1 to: 300

     1318 CGGGTTCCGACGCGGCAGAACGC 1341
          |||||||||||||||||||||||
      100 ArgValSerAspAlaAlaGluArg 107

seq_name: SwissProt_39:CYCR_RHOGE

seq_documentation_block:
```
ID   CYCR_RHOGE     STANDARD;      PRT;   366 AA.
AC   P51758;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   PHOTOSYNTHETIC REACTION CENTER CYTOCHROME C SUBUNIT PRECURSOR.
GN   PUFC.
OS   Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC   Rubrivivax.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=11144;
RX   MEDLINE=94132007; PubMed=8300574;
RA   Nagashima K.V.P., Matsuura K., Ohyama S., Shimada K.;
RT   "Primary structure and transcription of genes encoding B870 and
RT   photosynthetic reaction center apoproteins from Rubrivivax
RT   gelatinosus.";
RL   J. Biol. Chem. 269:2477-2484(1994).
CC   -!- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A
CC       TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO
CC       OXIDIZED PRIMARY ELECTRON DONOR.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC       ANCHOR (BY SIMILARITY).
CC   -!- PTM: BINDS FOUR HEME GROUPS PER MOLECULE.
CC   -!- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER
CC       CYTOCHROME C SUBUNITS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use  by non-profit institutions as long  as its content is  in  no  way
CC   modified and this statement is not removed.  Usage  by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D16822; BAA04102.1; -.
DR   HSSP; P07173; 4PRC.
DR   INTERPRO; IPR000345; -.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; 4.
KW   Electron transport; Photosynthesis; Reaction center; Heme;
KW   Membrane; Lipoprotein; Duplication; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    366       PHOTOSYNTHETIC REACTION CENTER CYTOCHROME
FT                                C SUBUNIT.
FT   LIPID        23     23       N-ACYL DIGLYCERIDE (BY SIMILARITY).
FT   BINDING     107    107       HEME 1 (COVALENT) (BY SIMILARITY).
FT   BINDING     110    110       HEME 1 (COVALENT) (BY SIMILARITY).
FT   METAL       111    111       IRON 1 (HEME AXIAL LIGAND)
FT                                (BY SIMILARITY).
FT   BINDING     151    151       HEME 2 (COVALENT) (BY SIMILARITY).
FT   BINDING     154    154       HEME 2 (COVALENT) (BY SIMILARITY).
FT   METAL       155    155       IRON 2 (HEME AXIAL LIGAND)
FT                                (BY SIMILARITY).
FT   BINDING     249    249       HEME 3 (COVALENT) (BY SIMILARITY).
FT   BINDING     252    252       HEME 3 (COVALENT) (BY SIMILARITY).
FT   METAL       253    253       IRON 3 (HEME AXIAL LIGAND)
FT                                (BY SIMILARITY).
FT   BINDING     309    309       HEME 4 (COVALENT) (BY SIMILARITY).
FT   BINDING     312    312       HEME 4 (COVALENT) (BY SIMILARITY).
FT   METAL       313    313       IRON 4 (HEME AXIAL LIGAND)
FT                                (BY SIMILARITY).
SQ   SEQUENCE  366 AA;  39195 MW;  2480B137663160031 CRC64;
```

alignment_scores:
          Quality: 8.00   Length: 8
          Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x CYCR_RHOGE

Align seg 1/1 to: CYCR_RHOGE from: 1 to: 366

      43 GCCGCCTTGCTGGCAGGCTGCGAA 66
         ||||||||||||||||||||||||
      17 AlaAlaLeuLeuAlaGlyCysGlu 24

seq_name: SwissProt_39:G154_MOUSE

seq_documentation_block:
```
ID   G154_MOUSE     STANDARD;      PRT;   429 AA.
```

```
AC   P50636;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   GAMETOGENESIS EXPRESSED PROTEIN GEG-154.
GN   GEG-154.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SWISS; TISSUE-OVARY;
RX   MEDLINE=9534162; PubMed=7615504;
RA   Lopez-Alanon D.M., del Mazo J.;
RT   "Cloning and characterization of genes expressed during gametogenesis
RT   of female and male mice.";
RL   J. Reprod. Fertil. 103:323-329(1995).
CC   -!- DEVELOPMENTAL STAGE: PREFERENTIALLY EXPRESSED IN BOTH SEXES DURING
CC       GAMETOGENESIS.
CC   ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; X71642; CAA50643.1; -.
KW   Transmembrane.
FT   TRANSMEM      39     59       POTENTIAL.
FT   TRANSMEM     154    174       POTENTIAL.
SQ   SEQUENCE   429 AA;  45213 MW;  7A3132693379C60F CRC64;

alignment_scores:
         Quality:        8.00                         Length:        8
         Ratio:          1.000                        Gaps:          0
Percent Similarity: 100.000         Percent Identity: 100.000

alignment_block:
US-09-388-090-3/rev x G154_MOUSE    ..

Align seg 1/1  to: G154_MOUSE  from: 1  to: 429

293 GGGTCGCTGTCGGCAAGCGGGTCG 270
    ||||||||||||||||||||||||
128 GlySerLeuSerAlaSerGlySer 135

seq_name: SwissProt_39:YIEO_HAEIN

seq_documentation_block:
ID   YIEO_HAEIN     STANDARD;      PRT;    463 AA.
AC   P44903;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   HYPOTHETICAL PROTEIN HI0852.
GN   HI0852.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20;
RX   MEDLINE=9350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
```

```
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (PROBABLE).
CC   -!- SIMILARITY: STRONG, TO E.COLI YIEO.
CC   -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC       AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC   ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; U32766; AAC22509.1; -.
DR   TIGR; HI0852; HI0852; -.
KW   Hypothetical protein; Transport; Transmembrane; Inner membrane.
FT   TRANSMEM      10     30       POTENTIAL.
FT   TRANSMEM      49     69       POTENTIAL.
FT   TRANSMEM      82    102       POTENTIAL.
FT   TRANSMEM     107    127       POTENTIAL.
FT   TRANSMEM     139    159       POTENTIAL.
FT   TRANSMEM     165    185       POTENTIAL.
FT   TRANSMEM     197    217       POTENTIAL.
FT   TRANSMEM     225    245       POTENTIAL.
FT   TRANSMEM     267    287       POTENTIAL.
FT   TRANSMEM     298    318       POTENTIAL.
FT   TRANSMEM     328    348       POTENTIAL.
FT   TRANSMEM     354    374       POTENTIAL.
FT   TRANSMEM     393    413       POTENTIAL.
FT   TRANSMEM     429    449       POTENTIAL.
SQ   SEQUENCE   463 AA;  50002 MW;  D725EFFD50F9F23F CRC64;

alignment_scores:
         Quality:        8.00                         Length:        8
         Ratio:          1.000                        Gaps:          0
Percent Similarity: 100.000         Percent Identity: 100.000

alignment_block:
US-09-388-090-3/rev x YIEO_HAEIN    ..

Align seg 1/1  to: YIEO_HAEIN  from: 1  to: 463

273 GTCGGAATCGGTTCGGCATTGCC 250
    |||||||||||||||||||||||
400 ValGlyIleGlyPheGlyIleAla 407

seq_name: SwissProt_39:SYE_BACST

seq_documentation_block:
ID   SYE_BACST      STANDARD;      PRT;    489 AA.
AC   P22249;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   GLUTAMYL-TRNA SYNTHETASE (EC 6.1.1.17) (GLUTAMATE--TRNA LIGASE)
DE   (GLURS).
GN   GLTX.
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91009314; PubMed=2120226;
RA   Breton R., Watson D., Yaguchi M., Lapointe J.;
RT   "Glutamyl-tRNA synthetases of Bacillus subtilis 168T and of Bacillus
RT   stearothermophilus. Cloning and sequencing of the gltX genes and
```

RT   comparison with other aminoacyl-tRNA synthetases.";
RL   J. Biol. Chem. 265:18248-18255(1990).
CC   -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + TRNA(GLU) = AMP +
CC   PYROPHOSPHATE + L-GLUTAMYL-TRNA(GLU).
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; M55072; AAA22494.1; -.
DR   PIR; B36090; SYBSES.
DR   HSSP; P27000; 1GLN.
DR   INTERPRO; IPR000924; -.
DR   INTERPRO; IPR001412; -.
DR   PFAM; PF00749; tRNA-synt_1c; 1.
DR   PRINTS; PR00987; TRNASYNTHGLU.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW   Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT   SIMILAR   11   21   "HIGH" REGION.
FT   SIMILAR   253   257   "KMKS" REGION.
FT   BINDING   256   256   ATP (BY SIMILARITY).
SQ   SEQUENCE   489 AA; 56183 MW; AB71266CB80A2E15 CRC64;

alignment_scores:
                Quality: 8.00   Length: 8
                Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3/rev x SYE_BACST
Align seg 1/1 to: SYE_BACST from: 1 to: 489

753 GAACAGCCGGCGGCCCCGGAATTGCC 730
    ||||||||||||||||||||||||||
121 GluGlnArgAlaAlaGlyIleAla 128

seq_name: SwissProt_39:DEGP_BARHE

seq_documentation_block:
ID   DEGP_BARHE   STANDARD;   PRT;   503 AA.
AC   P54925;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROBABLE PERIPLASMIC SERINE PROTEASE DO-LIKE PRECURSOR (EC 3.4.21.-)
DE   (HTRA ANTIGEN).
GN   DEGP OR HTRA.
OS   Bartonella henselae (Rochalimaea henselae).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bartonellaceae; Bartonella.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HOUSTON-1;
RX   MEDLINE=94299828; PubMed=8027347;
RA   Anderson B., Sims K., Regnery R., Robinson L., Schmidt M.J.,
RA   Goral S., Hager C., Edwards K.;
RT   "Detection of Rochalimaea henselae DNA in specimens from cat scratch
RT   disease patients by PCR.";
RL   J. Clin. Microbiol. 32:942-948(1994).
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C; ALSO KNOWN AS THE
CC   DEGP/DEGQ/DEGS FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.
CC   ---

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; L20127; AAA97430.1; -.
DR   MEROPS; S01.273; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00595; PDZ; 2.
DR   PFAM; PF00089; trypsin; 1.
DR   PRINTS; PR00834; PROTEASES2C.
KW   Hydrolase; Serine protease; Periplasmic; Signal.
FT   SIGNAL   1   18   POTENTIAL.
FT   CHAIN   19   503   PROBABLE PERIPLASMIC SERINE PROTEASE DO-
FT   LIKE.
FT   ACT_SITE   143   143   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   173   173   CHARGE RELAY SYSTEM (POTENTIAL).
FT   ACT_SITE   247   247   CHARGE RELAY SYSTEM (POTENTIAL).
SQ   SEQUENCE   503 AA; 54114 MW; 6CD9F4743282AF9E CRC64;

alignment_scores:
                Quality: 8.00   Length: 8
                Ratio: 1.000   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x DEGP_BARHE
Align seg 1/1 to: DEGP_BARHE from: 1 to: 503

640 GTGACCGCCGGCATCGTGTCCGCC 663
    ||||||||||||||||||||||||
215 ValThrAlaGlyIleValSerAla 222

seq_name: SwissProt_39:FXC1_HUMAN

seq_documentation_block:
ID   FXC1_HUMAN   STANDARD;   PRT;   553 AA.
AC   Q12948; Q9UP06;
DT   01-NOV-1997 (Rel. 35, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   FORKHEAD BOX PROTEIN C1 (FORKHEAD-RELATED PROTEIN FKHL7) (FORKHEAD-
DE   RELATED TRANSCRIPTION FACTOR 3) (FREAC-3).
GN   FOXC1 OR FKHL7 OR FREAC3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS S-112; M-126 AND L-131.
RX   MEDLINE=98282091; PubMed=9620769;
RA   Nishimura D.Y., Swiderski R.E., Alward W.L.M., Searby C.C.,
RA   Patil S.R., Bennet S.R., Kanis A.B., Gastier J.M., Stone E.M.,
RA   Sheffield V.C.;
RT   "The forkhead transcription factor gene FKHL7 is responsible for
RT   glaucoma phenotypes which map to 6p25.";
RL   Nat. Genet. 19:140-147(1998).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANTS ARA THR-82 AND MET-87.
RX   MEDLINE=99011252; PubMed=9792859;
RA   Mears A.J., Jordan T., Mirzayans F., Dubois S., Kume T., Parlee M.,
RA   Ritch R., Koop B., Kuo W.-L., Collins C., Marshall J., Gould D.B.,
RA   Pearce W., Carlsson P., Enerbaeck S., Morissette J., Bhattacharya S.,
RA   Hogan B., Raymond V., Walter M.A.;
RT   "Mutations of the forkhead/winged-helix gene, FKHL7, in patients with
RT   Axenfeld-Rieger anomaly.";
RL   Am. J. Hum. Genet. 63:1316-1328(1998).

```
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x FXC1_HUMAN

Align seg: 1/1  to: FXC1_HUMAN  from: 1  to: 553

        599 GGGTCGCTGCCATCGGCGCGCGCCCT 622
            ||||||||||||||||||||||||||
        275 GlySerLeuProSerAlaArgPro 282

seq_name: SwissProt_39:FXC1_MOUSE

seq_documentation_block:
ID   FXC1_MOUSE    STANDARD;         PRT;   553 AA.
AC   Q61572; Q61582; O88409;
DT   01-NOV-1997 (Rel. 35, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   FORKHEAD BOX PROTEIN C1 (FORKHEAD-RELATED PROTEIN FKHL7) (FORKHEAD-
DE   RELATED TRANSCRIPTION FACTOR 3) (FREAC-3) (TRANSCRIPTION FACTOR FKH-1)
DE   (MESODERM/MESENCHYME FORKHEAD 1) (MF-1).
GN   FOXC1 OR FKHL7 OR FREAC3 OR FKH1 OR MF1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98297351; PubMed=9635428;
RA   Kume T., Deng K.Y., Winfrey V., Gould D.B., Walter M.A., Hogan B.L.M.;
RT   "The forkhead/winged helix gene Mf1 is disrupted in the pleiotropic
RT   mouse mutation congenital hydrocephalus.";
RL   Cell 93:985-996(1998).
RN   [2]
RP   SEQUENCE OF 69-179 FROM N.A.
RC   STRAIN=129;
RX   MEDLINE=93361500; PubMed=7689224;
RA   Kaestner K.H., Lee K.H., Schloendorff J., Hiemisch H.,
RA   Monaghan A.P., Schuetz G.;
RT   "Six members of the mouse forkhead gene family are developmentally
RT   regulated.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:7628-7631(1993).
RN   [3]
RP   SEQUENCE OF 71-187 FROM N.A.
RX   MEDLINE=93387221; PubMed=8375339;
RA   Sasaki H., Hogan B.L.;
RT   "Differential expression of multiple fork head related genes during
RT   gastrulation and axial pattern formation in the mouse embryo.";
RL   Development 118:47-59(1993).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY EMBRYONIC TISSUES, INCLUDING
CC       PRECHONDROGENIC MESENCHYME, PERIOCULAR MESENCHYME, MENINGES,
CC       ENDOTHELIAL CELLS, AND KIDNEY.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.
CC   -!- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF045017; AAC24209.1; -.
DR   EMBL; L10406; AAA03159.1; -.
DR   EMBL; X71939; CAA50741.1; -.
DR   HSSP; Q63245; 2HFH.
DR   TRANSFAC; T02426; -.
DR   MGD; MGI:1347466; FOXC1.
DR   INTERPRO; IPR001766; -.
DR   PFAM; PF00250; Fork_head; 1.
DR   PRINTS; PR00053; FORKHEAD.
```

```
DR   PROSITE; PS00657; FORK_HEAD_1; 1.
DR   PROSITE; PS00658; FORK_HEAD_2; 1.
DR   PROSITE; PS50039; FORK_HEAD_3; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
FT   DOMAIN       28     33       POLY-ALA.
FT   DNA_BIND     77    168       FORK-HEAD.
FT   DOMAIN      169    173       POLY-ARG.
FT   DOMAIN      194    197       POLY-PRO.
FT   DOMAIN      264    274       POLY-SER.
FT   DOMAIN      375    386       POLY-GLY.
FT   DOMAIN      444    451       POLY-SER.
FT   DOMAIN      453    456       POLY-GLY.
FT   DOMAIN      486    496       POLY-ALA.
FT   CONFLICT    180    187       VKDKEEKG -> KKEITFIG (IN REF. 3).
SQ   SEQUENCE    553 AA;  56953 MW;  3CDD12F69CA4F217 CRC64;
```

alignment_scores:
```
          Quality:     8.00          Length:        8
            Ratio:    1.000            Gaps:        0
Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:
US-09-388-090-3 x FXC1_MOUSE   ..

Align seg 1/1 to: FXC1_MOUSE from: 1 to: 553

```
599 GGGTCGCTGCCATCGGCGCGCCCT 622
    ||||||||||||||||||||||||
277 GlySerLeuProSerAlaArgPro 284
```

```
OM of: US-09-388-090-3 to: SPTREMBL_15:*      out_format : pfs

Date: Mar 22, 2001 12:08 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame=n2p.model -DEV-xlp
 -Q=/cgn2.1/USPTO_spool/US09388090/runat.20032001_092951_25846/app_query.fasta_1.1474
 -DB=SPTREMBL_15 -OFMT=fastan -SUFFIX=oligo7.rspt -GAPOP=4.500
 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -FGAPEXT=60.000 -YGAPEXT=60.000
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
 -TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR_SCORE=quality
 -THR_MIN=7 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09388090 @CGN1_1_123 -NCPU=6 -ICPU=3
 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
 Query: US-09-388-090-3
 Query length: 1395
 Database: SPTREMBL_15:*
 Database sequences: 374700
 Database length: 117207915
 Search time (sec): 62.980000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:
Sequence               Strd Orig  ZScore   EScore    Len  Documentation
sp_bacteria:Q9JVT1      +  127.00 2346.04 3.4e-123   499  ! Q9JVT1 neisseria meningitidis (
sp_bacteria:Q44476      +   19.00  324.28 1.5e-10    473  ! Q44476 azotobacter vinelandii.
sp_bacteria:Q57155      +   19.00  324.27 1.5e-10    474  ! Q57155 pseudomonas aeruginosa.
sp_bacteria:Q9KJN6      +   13.00  211.59 0.0003     500  ! Q9k jn6 myxococcus xanthus. puta
sp_bacteria:035021      +   11.00  179.25 0.0371     224  ! 035021 bacillus subtilis. yvtb
sp_bacteria:P72780      +   11.00  175.66 0.0335     394  ! P72780 synechocystis sp. (strai
sp_bacteria:006670      +   11.00  175.61 0.0334     397  ! 006670 streptococcus pneumoniae
sp_bacteria:083752      +   11.00  175.59 0.0334     398  ! 083752 treponema pallidum. peri
sp_plant:004674         +   11.00  175.59 0.0334     398  ! 004674 haematococcus pluvialis.
sp_bacteria:P73940      +   11.00  175.31 0.0331     416  ! P73940 synechocystis sp. (strai
sp_plant:Q9LK85         +   11.00  174.97 0.0328     439  ! Q9lk85 arabidopsis thaliana (mc
sp_bacteria:Q9RTK4      +   11.00  174.84 0.0328     441  ! Q9rtk4 deinococcus radiodurans.
sp_plant:Q9LU10         +   11.00  174.83 0.0327     448  ! Q9lu10 arabidopsis thaliana (mc
sp_bacteria:034358      +   11.00  174.77 0.0327     449  ! 034358 bacillus subtilis. ykda
sp_bacteria:P73354      +   11.00  174.77 0.0326     452  ! P73354 synechocystis sp. (strai
sp_bacteria:067436      +   11.00  174.70 0.0326     453  ! 067436 aquifex aeolicus. peripl
sp_bacteria:Q9R9I1      +   11.00  174.49 0.0326     458  ! Q9r9i1 bacillus subtilis. htra
sp_bacteria:Q9K1W4      +   11.00  174.30 0.0325     459  ! Q9k1w4 chlamydia pneumoniae (he
sp_bacteria:Q9QD7       +   11.00  174.30 0.0322     488  ! Q9jqd7 chlamydia pneumoniae (ch
sp_bacteria:Q9LP97      +   11.00  174.18 0.0321     488  ! Q9p197 chlamydia muridarum. ser
sp_plant:004481         +   11.00  174.18 0.0321     497  ! 004481 arabidopsis thaliana (mc
sp_plant:Q9SH20         +   10.00  156.09 0.3605     450  ! 009shz0 arabidopsis thaliana (mc
sp_plant:Q9SH20         +   10.00  155.19 0.3512     518  ! 082261 arabidopsis thaliana (mc
sp_bacteria:086761      +   10.00  153.93 0.3386     632  ! 086761 streptomyces coelicolor.
sp_bacteria:Q9S2R5      -   11.00  140.05 4.30       295  ! Q9s2r5 streptomyces coelicolor.
sp_bacteria:Q51374      +    9.00  138.75 4.14       362  ! Q51374 pseudomonas aeruginosa.
sp_plant:Q9K5R6         +    9.00  138.29 4.09       389  ! Q9k5r6 bacillus halodurans. ser
sp_plant:Q9LA06         +    9.00  138.02 4.05       406  ! Q9la06 lactococcus lactis. seri
sp_plant:Q9W41          +    9.00  137.99 4.04       408  ! Q9w41 lactobacillus helveticus
sp_invertebrate:Q9VFJ3  +    9.00  137.77 4.03       412  ! Q9z4h7 drosophila melanogaste
sp_bacteria:Q9QD7       +    9.00  137.34 3.98       422  ! Q9vfj3 drosophila melanogaste
sp_bacteria:Q925G6      +    9.00  137.17 3.96       452  ! Q92596 mycobacterium leprae. pu
sp_bacteria:053896      +    9.00  137.03 3.94       464  ! 053896 mycobacterium tuberculos
sp_bacteria:044596      +    9.00  136.52 3.88       474  ! 044596 brucella abortus. htra-l
sp_plant:Q9PBA3         +    9.00  134.37 3.65       514  ! Q9pba3 xylella fastidiosa. peri
sp_plant:080829         +    9.00  134.23 3.63       720  ! 080829 arabidopsis virilis (fr
sp_invertebrate:096641  +    9.00  133.78 3.59       790  ! 096641 drosophila virilis (fr
sp_plant:080394         -    9.00  133.53 3.56       822  ! 080394 mesembryanthemum crystal. 1123
sp_invertebrate:000901  -                            822  ! 000901 leishmania major. 1123
```

```
sp_plant:Q9ZPU5         -  9.00  130.45  3.26  ! Q9zpu5 arabidopsis thalian.    1335
sp_plant:Q9SH77         -  9.00  130.35  3.25  ! Q9sh77 arabidopsis thalian.    1356
sp_invertebrate:Q9M8I9  +  9.00  128.86  3.11  ! Q9m8i9 arabidopsis thalian     1713
sp_bacteria:Q9LR9       +  9.00  128.20  3.05  ! Q9ndy7 leishmania major.       1899
sp_invertebrate:Q9VFS8  +  8.00  124.23  51.62 ! Q91946 pseudomonas putida.      187
sp_bacteria:052255      +  8.00  123.20  50.10 ! Q9vfs8 drosophila melanog.      220
sp_bacteria:Q9S2P8      +  8.00  123.11  49.97 ! 052255 pseudomonas putida.      223
sp_human:095517         +  8.00  121.62  47.86 ! Q9s2p8 streptomyces coelico     282
sp_invertebrate:Q93727  -  8.00  121.44  47.61 ! 095517 homo sapiens (human)     290
sp_archea:Q9V0V5        +  8.00  121.26  47.38 ! Q93727 caenorhabditis elec      298
sp_bacteria:Q9S0K5      +  8.00  121.22  47.33 ! Q9v0v5 pyrococcus abyssi. t      300
sp_archea:Q9YDE9        +  8.00  121.05  47.09 ! Q9s0k5 shewanella violacea.      308
sp_bacteria:Q9ZE02      +  8.00  120.99  47.00 ! Q9yde9 aeropyrum pernix. hy      311
sp_bacteria:P95617      +  8.00  120.08  45.78 ! Q9ze02 rickettsia prowazeki      359
sp_bacteria:Q9JPB7      +  8.00  119.96  45.62 ! P95617 rhodocyclus gelatino      366
sp_bacteria:Q46094      +  8.00  119.92  45.57 ! Q9jpb7 rhodocyclus gelatino      368
sp_bacteria:031388      +  8.00  119.87  45.50 ! Q46094 campylobacter jejuni      371
sp_plant:Q9ZPG7         +  8.00  119.52  45.04 ! 031388 bradyrhizobium japon      392
sp_bacteria:Q9K6X6      +  8.00  119.42  44.92 ! Q9zpg7 arabidopsis thaliana     398
sp_human:Q9Y4Y1         +  8.00  119.34  44.82 ! Q9k6x6 bacillus halodurans.     403
sp_bacteria:025663      -  8.00  118.74  44.04 ! Q9y4y1 homo sapiens (human)     443
sp_invertebrate:044021  -  8.00  118.60  43.86 ! 025663 helicobacter pylori (    453
sp_human:Q9NV58         +  8.00  118.54  43.79 ! 044021 plasmodium falcipar      457
sp_bacteria:006439      +  8.00  118.50  43.74 ! Q9nv58 homo sapiens (human)     460
sp_bacteria:046120      +  8.00  118.34  43.53 ! 006439 rhodobacter capsulat     472
sp_bacteria:Q9PN69      +  8.00  118.34  43.53 ! 046120 campylobacter jejuni     472
sp_bacteria:Q9ZM18      +  8.00  118.28  43.46 ! Q9pn69 campylobacter jejuni     476
sp_bacteria:Q9PGL3      +  8.00  118.22  43.38 ! Q9zm18 helicobacter pylori j    481
sp_plant:Q9SRL6         +  8.00  117.92  43.01 ! Q9pgl3 xylella fastidiosa. h    504
sp_human:Q9NUE5         +  8.00  117.72  42.76 ! Q9vfm0 drosophila melanogas     520
sp_rodent:Q90WR9        +  8.00  117.33  42.28 ! Q9srl6 arabidopsis thaliana     553
sp_fungi:059954         +  8.00  117.33  42.28 ! Q9nue5 homo sapiens (human)     553
sp_vertebrate:093399    +  8.00  117.24  42.17 ! Q90wr9 mus musculus (mouse)     560
sp_invertebrate:Q9X8J0  -  8.00  117.05  41.94 ! 059954 emericella nidulans (    561
sp_rodent:Q9QXA2        +  8.00  117.05  41.94 ! 093399 xenopus laevis (afric    578
sp_plant:Q9SA69         +  8.00  116.71  41.53 ! Q9x8j0 streptomyces coelico     578
sp_vertebrate:042182    -  8.00  116.41  41.23 ! Q9w324 drosophila melanoga      610
sp_fungi:Q06825         +  8.00  116.01  40.69 ! Q9qxa2 mus musculus (mouse)     634
sp_plant:Q9XE97         +  8.00  115.51  40.12 ! Q9sa69 arabidopsis thaliana     681
sp_invertebrate:Q9V515  +  8.00  115.43  40.02 ! Q06825 saccharomyces cerevis    736
sp_rodent:P96323        +  8.00  115.16  39.71 ! Q9v515 drosophila melanoga      746
sp_rodent:Q90UJ5        +  8.00  114.93  39.44 ! Q9xe97 arabidopsis thaliana     778
sp_bacteria:087943      +  8.00  114.67  39.15 ! P96323 escherichia coli. put    807
sp_plant:Q9LYT4         +  8.00  114.51  38.97 ! Q90uj5 mus musculus (mouse)     840
sp_virus:Q91MX8         +  8.00  112.89  37.18 ! 087943 thauera aromatica. be    861
sp_fungi:Q03016         +  8.00  112.63  36.91 ! Q9lyt4 arabidopsis thaliana    1112
sp_plant:Q9MIF5         +  8.00  112.01  36.25 ! Q91mx8 cercopithicine herpe    1157
sp_invertebrate:Q17163  +  8.00  111.59  35.82 ! Q03016 saccharomyces cerevi    1276
sp_fungi:Q9Y8A2         +  8.00  110.92  35.13 ! Q9mlf5 arabidopsis thaliana    1363
sp_rodent:Q9WTS1        +  8.00  109.81  34.02 ! Q9s2y9 streptomyces coelico    1514
sp_rodent:Q9JLR6        +  7.00  117.47  31.79 ! Q17163 brugia malayi. a2       1802
sp_virus:Q90088         +  7.00  117.36  802.77 ! Q9y8a2 gibberella fujikuroi     2607
sp_virus:Q9YN14         +  7.00  116.17  775.47 ! Q9wts1 rattus norvegicus (rat    28
sp_bacteria:Q9K2L1      +  7.00  116.17  775.47 ! Q9jlr6 rattus norvegicus (rat    29
sp_virus:Q84308         +  7.00  115.16  753.22 ! Q90088 human papillomavirus.     35
sp_bacteria:Q9RR49      +  7.00  114.41  743.50 ! Q9yn14 human papillomavirus t    35
sp_bacteria:Q9KMH3      +  7.00  113.41  716.01 ! Q9k2l1 vibrio cholerae. hypot    41
sp_bacteria:Q9KME8      +  7.00  112.85  704.44 ! Q84308 azotobacter chroococcu    44
sp_bacteria:Q9KMC7      +  7.00  112.85  704.44 ! Q9rr49 vibrio cholerae. hypot    54
sp_bacteria:Q9KMF3      +  7.00  112.74  702.26 ! Q9kmh3 vibrio cholerae. hypot    59
sp_plant:Q9LF23         +  7.00  111.67  680.85 ! Q9kme8 vibrio cholerae. hypot    59
sp_plant:Q9XJS5         +  7.00  111.67  680.85 ! Q9kmc7 vibrio cholerae. hypot    59
sp_bacteria:Q9JXC0      +  7.00  111.40  675.68 ! Q9kmf3 vibrio cholerae. hypot    60
sp_virus:072946         +  7.00  110.60  660.11 ! Q9jwm2 neisseria meningitidis   71
sp_bacteria:Q9RUX4      +  7.00  110.37  655.86 ! Q9lf23 arabidopsis thaliana     71
sp_bacteria:Q9Z384      +  7.00  110.30  654.48 ! Q9xjs5 bacteriophage pm2. gp    74
sp_phage:Q38613         +  7.00  109.95  647.86 ! Q9jxc0 neisseria meningitidis   80
sp_bacteria:Q54156                              ! 072946 hepatitis c virus. gly    84
                                                ! Q9rux4 deinococcus radioduran   87
                                                ! Q9z384 actinobacillus actinom   88
                                                ! Q38613 bacteriophage tuc2009.   88
                                                ! Q54156 shigella flexneri. sal   93
```

Fri Mar 23 07:26:36 2001

us-09-388-090-3.oligo7.rspt

Entry	±	col	score1	score2	rank	Accession / description
sp_virus:Q91294	+	7.00	109.88	646.59	94	! Q91294 human immunodeficiency vi
sp_virus:090727	+	7.00	109.68	642.86	97	! 090727 human papillomavirus type
sp_virus:Q91GT5	+	7.00	109.68	642.86	97	! Q9igt5 porcine adenovirus 3. 97
sp_vertebrate:Q9PSD5	+	7.00	109.43	638.10	101	! Q9psd5 gallus gallus (chicken)
sp_bacteria:Q51757	-	7.00	109.24	634.67	104	Q51757 pseudomonas fluorescen
sp_bacteria:O26041	-	7.00	109.00	630.28	108	O26041 helicobacter pylori (cam
sp_bacteria:P94670	-	7.00	109.00	630.28	108	P94670 fremyella diplosiphon (c
sp_invertebrate:Q9VAA6	+	7.00	108.82	627.11	111	! Q9vaa6 drosophila melanogaste
sp_bacteria:Q9QRC2	-	7.00	108.77	626.08	112	Q9qrc2 human immunodeficiency v
sp_bacteria:069507	-	7.00	108.54	622.05	116	069507 mycobacterium leprae. hy
sp_bacteria:Q9KTG0	+	7.00	108.49	621.07	116	Q9ktg0 vibrio cholerae. hypothe
sp_archea:Q9Y9V3	-	7.00	108.38	619.13	117	Q9y9v3 aeropyrum pernix. hypoth
sp_bacteria:Q9X8E5	-	7.00	108.22	616.30	119	Q9x8e5 streptomyces coelicolor.
sp_archea:Q9YAB3	-	7.00	107.97	614.46	122	Q9yab3 aeropyrum pernix. hypoth
sp_archea:Q9Y912	-	7.00	107.97	611.77	124	Q9y912 aeropyrum pernix. hypoth
sp_virus:009503	-	7.00	107.77	608.29	127	009503 garlic virus b. nucleic
sp_plant:Q9SUX3	-	7.00	107.67	606.59	131	Q9sux3 arabidopsis thaliana (mc
sp_archea:031429	-	7.00	107.44	603.30	137	031429 bacillus subtilis. ybdd
sp_archea:Q9YE28	-	7.00	107.44	602.49	138	Q9ye28 aeropyrum pernix. hypot
sp_bacteria:P73805	-	7.00	107.35	600.90	140	P73805 synechocystis sp. (strai
sp_virus:Q9J9G8	-	7.00	107.26	599.33	142	Q9j9g8 coxsackievirus b4. polyp
sp_bacteria:066357	-	7.00	107.12	597.03	145	066357 unidentified nitrogen-fi
sp_mammal:Q28568	-	7.00	107.12	597.03	145	Q28568 ovis aries (sheep). earl
sp_plant:004166	-	7.00	107.12	597.03	145	004166 physcomitrella patens (m
sp_virus:073443	-	7.00	107.12	597.03	145	073443 human papillomavirus. la
sp_bacteria:Q43956	-	7.00	106.95	594.05	149	Q43956 azotobacter chroococcum
sp_bacteria:Q9X549	-	7.00	106.95	594.05	149	Q9x549 corynebacterium glutamic
sp_bacteria:Q9L266	-	7.00	106.78	591.16	153	Q9L266 streptomyces coelicolor.
sp_invertebrate:Q17550	+	7.00	106.78	591.16	153	Q17550 caenorhabditis elegans
sp_mammal:Q9XS44	-	7.00	106.78	591.16	153	Q9xs44 tayassu tajacu (collared
sp_virus:081071	-	7.00	106.78	591.16	153	081071 human papillomavirus. ma
sp_virus:Q84228	-	7.00	106.78	591.16	153	Q84228 human papillomavirus. 11
sp_virus:081073	-	7.00	106.78	591.16	153	081073 human papillomavirus. ma
sp_plant:049945	-	7.00	106.66	589.05	156	049945 solanum tuberosum (potat
sp_mammal:097659	-	7.00	106.36	584.31	163	097659 oryctolagus cuniculus (r
sp_invertebrate:Q9N5V2	+	7.00	106.08	579.19	171	! Q9n5v2 caenorhabditis elegans
sp_rodent:Q9ZLY2	-	7.00	105.96	577.95	173	Q9zly2 cavia porcellus (guinea
sp_invertebrate:Q9N474	+	7.00	105.96	577.34	174	! Q9n474 caenorhabditis elegans
sp_virus:Q9QEK8	-	7.00	105.93	576.73	175	Q9qek8 canine distemper virus.
sp_vertebrate:P70069	-	7.00	105.82	576.12	176	P70069 gallus gallus (chicken)
sp_human:Q9P1R8	-	7.00	105.82	574.93	178	Q9p1r8 homo sapiens (human). ho
sp_bacteria:084363	-	7.00	105.75	573.75	180	084363 chlamydia trachomatis. h
sp_human:Q9UF32	-	7.00	105.75	573.16	181	Q9uf32 homo sapiens (human). hy
sp_fungi:Q9USE4	-	7.00	105.71	573.16	181	Q9use4 schizosaccharomyces pomb
sp_invertebrate:Q9VSB3	-	7.00	105.64	572.05	183	! Q9vsb3 drosophila melanogaste
sp_plant:Q9M054	-	7.00	105.61	571.43	184	Q9m054 arabidopsis thaliana (mc
sp_virus:Q9IGT3	-	7.00	105.51	569.73	187	Q9igt3 porcine adenovirus 3. 18
sp_bacteria:Q07394	+	7.00	105.51	569.73	187	Q07394 rhodobacter capsulatus (
sp_bacteria:Q9KYR3	+	7.00	105.44	568.62	189	Q9kyr3 streptomyces coelicolor.
sp_invertebrate:Q9VD16	+	7.00	105.44	568.62	189	! Q9vd16 drosophila melanogaste
sp_organelle:Q9MF32	+	7.00	105.44	568.62	189	Q9mf32 brachyderes grisescens.
sp_organelle:Q9MF30	+	7.00	105.44	568.07	190	Q9mf30 brachyderes pubescens. c
sp_mammal:002834	+	7.00	105.40	568.07	190	002834 sus scrofa (pig). adipoc
sp_organelle:Q9MF29	-	7.00	105.40	568.07	190	Q9mf29 brachyderes rugatus scul
sp_organelle:Q9MF28	-	7.00	105.40	568.07	190	Q9mf28 brachyderes rugatus scul
sp_organelle:Q9MF26	-	7.00	105.40	568.07	190	Q9mf26 brachyderes rugatus scul
sp_organelle:Q9MF25	-	7.00	105.40	568.07	190	Q9mf25 brachyderes rugatus scul
sp_organelle:Q9MF24	-	7.00	105.40	568.07	190	Q9mf24 brachyderes rugatus scul
sp_organelle:Q9MF23	-	7.00	105.40	568.07	190	Q9mf23 brachyderes rugatus scul
sp_organelle:Q9MF22	-	7.00	105.40	568.07	190	Q9mf22 brachyderes rugatus scul
sp_organelle:Q9MF21	-	7.00	105.40	568.07	190	Q9mf21 brachyderes rugatus scul
sp_organelle:Q9MF20	-	7.00	105.40	568.07	190	Q9mf20 brachyderes rugatus scul
sp_organelle:Q9MF19	-	7.00	105.40	568.07	190	Q9mf19 brachyderes rugatus scul
sp_organelle:Q9MF18	-	7.00	105.40	568.07	190	Q9mf18 brachyderes rugatus scul
sp_organelle:Q9MF17	-	7.00	105.40	568.07	190	Q9mf17 brachyderes rugatus scul
sp_organelle:Q9MF16	-	7.00	105.40	568.07	190	Q9mf16 brachyderes rugatus scul
sp_organelle:Q9MF14	-	7.00	105.40	568.07	190	Q9mf14 brachyderes rugatus scul
sp_organelle:Q9MF13	-	7.00	105.40	568.07	190	Q9mf13 brachyderes rugatus ruga
sp_organelle:Q9MF12	-	7.00	105.40	568.07	190	Q9mf12 brachyderes rugatus ruga
sp_organelle:Q9MF11	-	7.00	105.40	568.07	190	Q9mf11 brachyderes rugatus ruga
sp_organelle:Q9MF10	-	7.00	105.40	568.07	190	Q9mf10 brachyderes rugatus r
sp_organelle:Q9MF09	-	7.00	105.40	568.07	190	Q9mf09 brachyderes rugatus r
sp_organelle:Q9MF08	-	7.00	105.40	568.07	190	Q9mf08 brachyderes rugatus r
sp_organelle:Q9MF07	-	7.00	105.40	568.07	190	Q9mf07 brachyderes rugatus r
sp_organelle:Q9MF06	-	7.00	105.40	568.07	190	Q9mf06 brachyderes rugatus r
sp_organelle:Q9MF05	-	7.00	105.40	568.07	190	Q9mf05 brachyderes rugatus r
sp_organelle:Q9MF04	-	7.00	105.40	568.07	190	Q9mf04 brachyderes rugatus r
sp_organelle:Q9MF03	-	7.00	105.40	568.07	190	Q9mf03 brachyderes rugatus h
sp_organelle:Q9MF02	-	7.00	105.40	568.07	190	Q9mf02 brachyderes rugatus h
sp_organelle:Q9MF01	-	7.00	105.40	568.07	190	Q9mf01 brachyderes rugatus h
sp_organelle:Q9MF00	-	7.00	105.40	568.07	190	Q9mf00 brachyderes rugatus h
sp_organelle:Q9MEZ9	-	7.00	105.40	568.07	190	Q9mez9 brachyderes rugatus h
sp_organelle:Q9MEZ8	-	7.00	105.40	568.07	190	Q9mez8 brachyderes rugatus h
sp_organelle:Q9MEZ7	-	7.00	105.40	568.07	190	Q9mez7 brachyderes rugatus h
sp_organelle:Q9MDU5	-	7.00	105.40	568.07	190	Q9mdu5 brachyderes rugatus c
sp_organelle:Q9MDT3	-	7.00	105.40	568.07	190	Q9mdt3 brachyderes rugatus a
sp_invertebrate:Q9VGN9	+	7.00	105.37	567.52	191	! Q9vgn9 drosophila melanoga
sp_vertebrate:Q9YB35	-	7.00	105.31	566.97	192	Q9yb35 aeropyrum pernix. hyp
sp_vertebrate:Q9J9F8	-	7.00	105.31	566.43	193	Q9j9f8 dendroica adelaidae.
sp_vertebrate:Q9J9F7	-	7.00	105.31	566.43	193	Q9j9f7 basileuterus culicivo
sp_archea:029261	-	7.00	105.27	565.89	194	029261 archaeoglobus fulgidu
sp_archea:058972	-	7.00	105.27	565.89	194	058972 pyrococcus horikoshii
sp_bacteria:Q45576	-	7.00	105.27	565.89	194	Q45576 bacillus stearothermo
sp_bacteria:Q9RS76	-	7.00	105.27	565.89	194	Q9rs76 deinococcus radiodura
sp_plant:Q9ZPT2	-	7.00	105.21	564.83	196	Q9zpt2 arabidopsis thaliana
sp_bacteria:P95037	-	7.00	105.14	563.77	198	P95037 mycobacterium tubercu
sp_human:075983	-	7.00	105.14	563.77	198	075983 homo sapiens (human)
sp_bacteria:Q9KSL6	-	7.00	105.08	562.73	200	Q9ksl6 vibrio cholerae. alph
sp_plant:Q9LWE5	-	7.00	105.02	561.70	202	Q9lwe5 oryza sativa (rice)
sp_bacteria:Q9JXE1	-	7.00	105.02	561.70	202	Q9jxe1 neisseria meningitidi
sp_bacteria:Q9JWJ9	-	7.00	105.02	561.70	202	Q9jwj9 neisseria meningitidi
sp_archea:052008	-	7.00	104.95	560.68	204	052008 halobacterium sp. orf
sp_bacteria:Q9ZA40	-	7.00	104.95	560.68	204	Q9za40 klebsiella pneumoniae
sp_bacteria:Q9L3G5	-	7.00	104.92	560.18	205	Q9x6b5 yersinia pestis. dime
sp_bacteria:Q9X6B5	-	7.00	104.83	558.69	208	Q9z575 streptomyces coelicol
sp_bacteria:Q9Z575	-	7.00	104.71	556.73	212	014811 homo sapiens (human)
sp_human:014811	-	7.00	104.71	556.73	212	Q9wv43 mus musculus (mouse)
sp_rodent:Q9WV43	-	7.00	104.68	556.25	213	069361 feline herpesvirus (f
sp_virus:069361	-	7.00	104.59	554.84	216	Q9pup6 coltus striatus. oocy
sp_vertebrate:Q9PUP6	-	7.00	104.56	554.35	217	089116 mus musculus (mouse)
sp_rodent:089116	-	7.00	104.56	554.35	217	Q9ji52 rattus norvegicus (ra
sp_rodent:Q9JI52	-	7.00	104.47	552.95	220	Q9za40 streptomyces violaceo
sp_bacteria:Q9ZA40	-	7.00	104.41	552.03	222	Q9k3k8 streptomyces coelicol
sp_bacteria:Q9K3K8	-	7.00	104.41	552.03	222	Q9jqq9 neisseria meningitidi
sp_bacteria:Q9JQQ9	-	7.00	104.36	551.12	224	086695 streptomyces coelicol
sp_bacteria:086695	-	7.00	104.36	551.12	224	Q9ji51 rattus norvegicus (ra
sp_rodent:Q9JI51	-	7.00	104.22	550.22	226	Q9sh12 arabidopsis thaliana
sp_plant:Q9SH12	-	7.00	104.22	548.89	229	P96689 bacillus subtilis. yd
sp_bacteria:P96689	-	7.00	104.22	548.89	229	084996 streptococcus pneumon
sp_bacteria:084996	-	7.00	104.22	548.89	229	Q9s6u9 streptococcus pneumon
sp_bacteria:Q9S6U9	-	7.00	104.19	548.45	230	Q9jql5 neisseria meningitidi
sp_bacteria:Q9JQL5	-	7.00	104.08	546.71	234	- Q9nnr2 leishmania major.
sp_invertebrate:Q9NNR2	-	7.00	104.05	546.28	235	Q9ns28 homo sapiens (human)
sp_human:Q9NS28	-	7.00	104.00	545.43	237	026387 methanobacterium ther
sp_archea:026387	-	7.00	104.00	545.43	237	Q9rsp6 deinococcus radiodura
sp_bacteria:Q9RSP6	-	7.00	103.95	544.59	239	023959 glycine max (soybean)
sp_plant:023959	-	7.00	103.92	544.17	240	Q9xhx6 oryza sativa (rice)
sp_plant:Q9XHX6	-	7.00	103.89	543.75	241	Q9xeq4 sorghum bicolor (sorg
sp_plant:Q9XEQ4	-	7.00	103.84	542.93	243	Q9ydz4 aeropyrum pernix. 243
sp_archea:Q9YDZ4	-	7.00	103.84	542.93	243	Q9i9u0 lagopus leucurus (whi
sp_vertebrate:Q9I9U0	-	7.00	103.84	542.93	243	Q9i9t9 lagopus leucurus (whi
sp_vertebrate:Q9I9T9	-	7.00	103.79	542.11	245	075062 homo sapiens (human)
sp_human:075062	-	7.00	103.79	542.11	245	Q9lh11 arabidopsis thaliana
sp_plant:Q9LH11	-	7.00	103.76	541.70	246	- Q22895 caenorhabditis eleq
sp_invertebrate:Q22895	-	7.00	103.74	541.30	247	P95555 pseudomonas stutzeri
sp_bacteria:P95555	-	7.00	103.74	541.30	247	Q9rd63 streptomyces coelicol
sp_bacteria:Q9RD63	-	7.00	103.74	541.30	247	Q9svc5 arabidopsis thaliana
sp_plant:Q9SVC5	-	7.00	103.92	543.75	240	Q9xhx6
sp_bacteria:Q9PRB7	-	7.00	103.71	540.90	248	Q9prb7 urcaplasma urealyticu
sp_bacteria:Q9PKF7	-	7.00	103.71	540.90	248	Q9pkf7 chlamydia muridarum.
sp_invertebrate:Q9VKG5	-	7.00	103.71	540.90	248	- Q9vkg5 drosophila melanoga
sp_bacteria:Q9RXJ6	-	7.00	103.68	540.50	249	Q9rxj6 deinococcus radiodura
sp_organelle:Q9T584	-	7.00	103.68	540.50	249	Q9t584 porphyra purpurea. su
sp_bacteria:Q9RSP1	-	7.00	103.66	540.10	250	Q9rsp1 deinococcus radiodura

#	Accession	val1	val2		sp entry	description
329	Q9NP22	513.49	101.91	7.00	sp_human:Q9NP22	homo sapiens (human)
333	O31568	512.35	101.83	7.00	sp_bacteria:O31568	bacillus subtilis. y
333	Q9S849	512.35	101.83	7.00	sp_plant:Q9S849	arabidopsis thaliana
334	Q9r0d8	512.07	101.82	7.00	sp_rodent:Q9R0D8	mus musculus (mouse)
335	Q51587	511.78	101.80	7.00	sp_phage:Q51587	plectonema boryanum.
335	Q9ks12	511.22	101.76	7.00	sp_bacteria:Q9KSL2	vibrio cholerae. vit
337	Q9zgh3	511.22	101.76	7.00	sp_bacteria:Q92GH3	streptomyces venezue
337	Q9uns5	511.22	101.76	7.00	sp_human:Q9UNS5	homo sapiens (human)
337	Q9sv48			7.00	sp_plant:Q9SV48	arabidopsis thaliana
339	! O45162	510.67	101.72	7.00	sp_invertebrate:O45162	caenorhabditis ele
341	P96358	510.12	101.68	7.00	sp_bacteria:P96358	mycobacterium tuberc
341	O95400	510.12	101.68	7.00	sp_human:O95400	homo sapiens (human)
343	Q9xbi7	509.57	101.65	7.00	sp_bacteria:Q9XBI7	bacillus cereus. 5-m
345	P70913	509.02	101.61	7.00	sp_bacteria:P70913	bacillus intermedius
345	Q9x1e2	509.02	101.61	7.00	sp_bacteria:Q9KIE2	streptomyces hygrosc
346	Q9z669	508.75	101.59	7.00	sp_bacteria:Q92669	zymomonas mobilis. h
347	Q9kzu4	508.48	101.57	7.00	sp_bacteria:Q9KZU4	streptomyces coelico
348	O30270	508.21	101.55	7.00	sp_archea:O30270	archaeoglobus fulgid
351	Q52013	507.41	101.50	7.00	sp_bacteria:Q52013	pseudomonas putida.
352	Q9kuf6	507.14	101.48	7.00	sp_bacteria:Q9KUF6	vibrio cholerae. pro
353	P95274	506.88	101.46	7.00	sp_bacteria:P95274	mycobacterium tuberc
357	! Q17682	505.83	101.39	7.00	sp_invertebrate:Q17682	caenorhabditis ele
357	Q9lia0	505.83	101.39	7.00	sp_plant:Q9LIA0	arabidopsis thaliana
358	Q9s270	505.57	101.37	7.00	sp_bacteria:Q9S270	streptomyces coelico
359	O53867	505.31	101.36	7.00	sp_bacteria:O53867	mycobacterium tuberc
359	O60759	505.05	101.34	7.00	sp_human:O60759	homo sapiens (human)
360	Q9x613	505.05	101.34	7.00	sp_bacteria:Q9X613	chlorobium tepidum.
360	Q9lhr4			7.00	sp_plant:Q9LHR4	arabidopsis thaliana
362	! Q27744	504.54	101.30	7.00	sp_invertebrate:Q27744	plasmodium falcipa
362	! Q9u6r2	504.54	101.30	7.00	sp_invertebrate:Q9U6R2	plasmodium falcipa
362	P93188	504.54	101.30	7.00	sp_plant:P93188	hordeum vulgare (bar
365	Q9q92	503.77	101.25	7.00	sp_bacteria:Q9Q92	rhizobium meliloti (
366	Q9kxn5	503.52	101.22	7.00	sp_bacteria:Q9KXN5	streptomyces coelico
367	Q9kps3	503.27	101.22	7.00	sp_bacteria:Q9KPS3	vibrio cholerae. thi
367	Q9se32	503.27	101.20	7.00	sp_plant:Q9SE32	oryza sativa (rice).
368	! Q9twa3	503.01	101.20	7.00	sp_invertebrate:Q9TWA3	plasmodium berghei
369	Q12593	502.76	101.18	7.00	sp_fungi:Q12593	cryphonectria parasit
369	Q92171	502.76	101.18	7.00	sp_vertebrate:Q92171	gallus gallus (chick
370	O29775	502.51	101.16	7.00	sp_archea:O29775	archaeoglobus fulgid
370	Q9rtu4	502.51	101.16	7.00	sp_bacteria:Q9RTU4	deinococcus radiodur
372	O97658	502.01	101.13	7.00	sp_mammal:O97658	bos taurus (bovine)
372	O33109	502.01	101.13	7.00	sp_bacteria:O33109	mycobacterium leprae
372	Q12121			7.00	sp_fungi:Q12121	saccharomyces cerevi
372	! Q9v4s3	502.01	101.13	7.00	sp_invertebrate:Q9V4S3	drosophila melanog
373	Q58859	501.77	101.11	7.00	sp_archea:Q58859	methanococcus jannas
373	O53253	501.77	101.11	7.00	sp_bacteria:O53253	mycobacterium tuberc
376	O85136	501.03	101.06	7.00	sp_bacteria:O85136	rhodobacter sphaeroid
378	Q01451	500.54	101.03	7.00	sp_bacteria:Q01451	haemophilus influenza
379	! Q9tyq1	500.30	101.01	7.00	sp_invertebrate:Q9TYQ1	caenorhabditis eleg
380	Q43568	499.57	100.99	7.00	sp_plant:Q43568	nicotiana tabacum (co
382	O69815	499.57	100.96	7.00	sp_bacteria:O69815	streptomyces coelicol
382	Q9rxp9	499.57	100.96	7.00	sp_bacteria:Q9RXP9	deinococcus radiodura
385	Q9x832	498.85	100.91	7.00	sp_bacteria:Q9X832	streptomyces coelicol
385	Q9pei6	498.85	100.91	7.00	sp_bacteria:Q9PEI6	xylella fastidiosa.
391	Q44032	497.44	100.81	7.00	sp_bacteria:Q44032	alcaligenes eutrophus
392	Q9sla1	497.20	100.80	7.00	sp_plant:Q9SLA1	arabidopsis thaliana
394	O30873	496.74	100.76	7.00	sp_bacteria:O30873	azotobacter chroococc
394	Q9x716	496.74	100.76	7.00	sp_bacteria:Q9X716	pseudomonas sp. p-hyd
394	Q02969	496.74	100.76	7.00	sp_fungi:Q02969	saccharomyces cerevis
394	Q14828	496.74	100.75	7.00	sp_human:Q14828	homo sapiens (human)
395	O47207	496.50	100.75	7.00	sp_bacteria:O47207	escherichia coli. fas
395	P77099	496.50	100.75	7.00	sp_bacteria:P77099	escherichia coli. iso
395	Q9r9t1	496.50	100.75	7.00	sp_bacteria:Q9R9T1	pseudomonas putida.
395	Q9r6l5	496.27	100.75	7.00	sp_bacteria:Q9R6L5	agrobacterium tumefac
396	O80871	496.27	100.73	7.00	sp_plant:O80871	arabidopsis thaliana
397	O53870	496.04	100.72	7.00	sp_bacteria:O53870	mycobacterium tubercu
397	! Q9nkf2	496.04	100.72	7.00	sp_invertebrate:Q9NKF2	drosophila melanoga
398	Q9x9x1	495.81	100.70	7.00	sp_vertebrate:Q9X9X1	colinus virginianus
400	Q9qy79	495.36	100.67	7.00	sp_rodent:Q9QY79	rattus norvegicus (ra
400	Q9jkm5	495.36	100.67	7.00	sp_rodent:Q9JKM5	rattus norvegicus (ra
401	! Q9nek9	495.13	100.65	7.00	sp_invertebrate:Q9NEK9	caenorhabditis eleg
402	O23054	494.91	100.64	7.00	sp_plant:O23054	arabidopsis thaliana
403	O42417	494.68	100.62	7.00	sp_vertebrate:O42417	gallus gallus (chicke

#	Accession	val1	val2		sign	sp entry	description
251	Q9i9x0	539.70	103.63	7.00	-	sp_vertebrate:Q9I9X0	colinus virginianus (bob
252	Q9zgg8	539.31	103.61	7.00	!	sp_bacteria:Q92GG8	heliobacillus mobilis. c
255	P94179	538.13	103.53	7.00	+	sp_bacteria:P94179	anabaena sp. (strain pcc
255	Q9xdx8	538.13	103.53	7.00	+	sp_bacteria:Q9XDX8	rhodopseudomonas palustr
256	Q38150	537.75	103.51	7.00	-	sp_phage:Q38150	bacteriophage spp1. phag
257	Q9n762	537.36	103.48	7.00	!	sp_invertebrate:Q9N762	leishmania major. poss
258	Q9wx18	536.98	103.46	7.00		sp_bacteria:Q9WX18	streptomyces coelicolor.
259	Q68105	536.59	103.43	7.00	+	sp_virus:Q68105	hawaii calicivirus. hypo
260	P96327	536.21	103.41	7.00	+	sp_bacteria:P96327	escherichia coli. putati
262	Q52602	535.46	103.36	7.00	+	sp_bacteria:Q52602	agrobacterium tumefacien
262	Q9rj31	535.46	103.36	7.00	+	sp_bacteria:Q9RJ31	streptomyces coelicolor.
262	! Q18146	535.46	103.36	7.00	-	sp_invertebrate:Q18146	caenorhabditis elegans
263	O05919	535.08	103.34	7.00		sp_bacteria:O05919	mycobacterium tuberculos
266	Q9961o	533.97	103.26	7.00	+	sp_bacteria:Q99619	homo sapiens (human). ch
266	Q9y915	533.97	103.26	7.00	+	sp_archea:Q9Y915	aeropyrum pernix. hypoth
271	Q9pib8	532.14	103.15	7.00		sp_plant:Q9XI34	arabidopsis thaliana (mc
271	O01644	532.14	103.15	7.00	!	sp_invertebrate:O01644	campylobacter jejuni. 2
271	082115	532.14	103.15	7.00		sp_invertebrate:O01644	drosophila grimshawi (
271	082115	532.14	103.15	7.00		sp_plant:082115	oryza sativa (rice). zin
276	Q9k4s9	530.35	103.03	7.00		sp_bacteria:Q9K4S9	pseudomonas putida. repl
281	Q9np23	528.61	102.92	7.00		sp_human:Q9NP23	homo sapiens (human). hy
281	Q9zse3	528.61	102.92	7.00		sp_plant:Q9ZSE3	euglena gracilis. fibril
282	Q9m4y9	528.23	102.89	7.00		sp_plant:Q9M4Y9	mesembryanthemum crystal
285	Q9rk77	527.23	102.83	7.00	+	sp_bacteria:Q9RK77	streptomyces coelicolor.
285	! Q22623	527.23	102.83	7.00	+	sp_invertebrate:Q22623	caenorhabditis elegans
285	P89001	526.89	102.80	7.00		sp_virus:P89001	mastomys natalensis papi
286	P87595	526.22	102.76	7.00		sp_virus:P87595	cowpox virus (cpv). 41kd
288	Q9k5n1	525.88	102.74	7.00		sp_bacteria:Q9K5N1	bacillus halodurans. sta
289	Q9jlu8	525.22	102.69	7.00		sp_rodent:Q9JLU8	cavia porcellus (guinea
291	Q9uu76	525.22	102.69	7.00		sp_bacteria:Q9UU76	neisseria meningitidis (
291	Q9np24	523.57	102.58	7.00		sp_human:Q9NP24	homo sapiens (human). ad
296	Q9xej3	523.25	102.56	7.00		sp_plant:Q9XEJ3	zea mays (maize). hypoth
297	! Q18574	523.25	102.56	7.00	-	sp_invertebrate:Q18574	caenorhabditis elegans
298	Q9k9h3	522.92	102.54	7.00		sp_bacteria:Q9K9H3	bacillus halodurans. 6-p
298	O43391	522.92	102.54	7.00		sp_human:O43391	homo sapiens (human). sy
298	O00173	522.92	102.54	7.00		sp_human:O00173	homo sapiens (human). me
298	O00560	522.92	102.54	7.00		sp_human:O00560	homo sapiens (human). me
299	Q9j4p3	522.60	102.52	7.00		sp_virus:Q9J4P3	human immunodeficiency v
302	Q9kk85	521.64	102.46	7.00		sp_bacteria:Q9KK85	brevibacterium linens. f
302	Q9x7m7	521.64	102.46	7.00		sp_bacteria:Q69914	streptomyces coelicolor.
302	Q9lojo	521.64	102.46	7.00		sp_bacteria:Q9X7M7	streptomyces coelicolor.
303	Q9n0u9	521.33	102.44	7.00		sp_bacteria:Q9LOJO	streptomyces coelicolor.
303	Q9y9d5	521.33	102.44	7.00		sp_mammal:Q9N0U9	ovis aries (sheep). amil
304	P71913	521.01	102.41	7.00		sp_archea:Q9Y9D5	aeropyrum pernix. hypoth
305	Q9y9j0	520.69	102.39	7.00		sp_bacteria:P71913	mycobacterium tuberculos
306	Q53929	520.38	102.37	7.00		sp_archea:Q9Y9J0	aeropyrum pernix. 305aa
307	O26926	520.07	102.35	7.00		sp_bacteria:Q53929	streptomyces coelicolor.
308	Q9ydy2	519.76	102.33	7.00		sp_archea:O26926	methanobacterium thermoa
309	O82583	519.45	102.31	7.00		sp_archea:Q9YDY2	aeropyrum pernix. hypoth
310	Q9uxj4	519.14	102.29	7.00		sp_plant:O82583	zantedeschia aethiopica
311	O67437	518.83	102.27	7.00		sp_archea:Q9UXJ4	sulfolobus solfataricus.
311	O33200	518.83	102.27	7.00		sp_bacteria:O67437	aquifex aeolicus. rna po
312	Q9rdh1	518.53	102.25	7.00		sp_bacteria:O33200	mycobacterium tuberculos
314	Q9rx37	517.92	102.21	7.00	!	sp_bacteria:Q9RDH1	streptomyces coelicolor.
314	O77442	517.92	102.21	7.00	+	sp_invertebrate:O77442	taenia saginata. heat
315	Q9kv12	517.61	102.19	7.00	-	sp_bacteria:Q9KV12	vibrio cholerae. trna de
317	Q9m8f9	517.01	102.15	7.00		sp_plant:Q9M8F9	arabidopsis thaliana (mc
320	Q9x832	516.12	102.09	7.00		sp_bacteria:Q9X832	streptomyces coelicolor.
320	Q9sjv0	516.12	102.09	7.00		sp_plant:Q9SJV0	arabidopsis thaliana (mc
321	O69128	515.82	102.07	7.00		sp_bacteria:O69128	burkholderia pseudomalle
321	Q9pfi2	515.82	102.07	7.00		sp_bacteria:Q9PFI2	xylella fastidiosa. coba
321	Q9jpc8	515.82	102.07	7.00		sp_bacteria:Q9JPC8	rhodocyclus gelatinosus.
321	Q99608	515.82	102.07	7.00		sp_human:Q99608	homo sapiens (human). ne
323	Q9klw1	515.23	102.03	7.00		sp_plant:Q9SEL7	arabidopsis thaliana d-3-pho
323	Q9p0y3	515.23	102.03	7.00		sp_bacteria:Q9KLW1	vibrio cholerae. d-3-pho
324	O32412	514.94	102.01	7.00		sp_human:Q9P0Y3	homo sapiens (human). se
324	Q9jqv2	514.94	102.01	7.00		sp_bacteria:O32412	rhodospirillum molischia
324	Q15630	514.94	102.01	7.00		sp_bacteria:Q9JQV2	neisseria meningitidis (
327	Q9x892	514.07	101.95	7.00	+	sp_human:Q15630	homo sapiens (human). tr
328	! Q20759	513.78	101.93	7.00	+	sp_bacteria:Q9X892	streptomyces coelicolor.
329	P96645	513.49	101.91	7.00	+	sp_invertebrate:Q20759	caenorhabditis elegans
329	P96645	513.49	101.91	7.00		sp_bacteria:P96645	bacillus subtilis. yddh

Left table

ID		7.00	%id	score	len	acc	description
sp_vertebrate:Q19K9	+	7.00	100.62	494.68	403	Q19k9	gallus gallus (chicken).
sp_rodent:Q9JIK3	+	7.00	100.57	494.00	406	Q9jik3	rattus norvegicus (rat).
sp_plant:Q9M2M8	-	7.00	100.53	493.33	409	Q9m2m8	arabidopsis thaliana (mc
sp_plant:Q9M9Z1	-	7.00	100.48	492.67	412	Q9m9z1	arabidopsis thaliana (mc
sp_bacteria:Q9RYF0	+	7.00	100.43	492.01	415	Q9ryf0	deinococcus radiodurans.
sp_plant:Q9SJH8	+	7.00	100.43	492.01	415	Q9sjh8	arabidopsis thaliana (mc
sp_virus:066626	-	7.00	100.43	492.01	415	066626	equine herpesvirus 2. or
sp_invertebrate:Q17495	+	7.00	100.40	491.58	417	!Q17495	caenorhabditis elegans
sp_archea:Q9UY07	+	7.00	100.40	491.15	419	Q9uy07	pyrococcus abyssi. gluco
sp_invertebrate:Q24548	+	7.00	100.37	491.15	419	!Q24548	drosophila melanogaste
sp_plant:Q9LUT4	+	7.00	100.37	491.15	419	Q9lut4	arabidopsis thaliana (mc
sp_human:Q9NP21	-	7.00	100.37	490.72	421	Q9np21	homo sapiens (human). (mc
sp_bacteria:051572	+	7.00	100.31	490.29	423	051572	borrelia burgdorferi (ly
sp_human:Q9P0Y4	+	7.00	100.31	490.29	423	Q9p0y4	homo sapiens (human). se
sp_bacteria:Q9SM16	-	7.00	100.30	490.07	424	Q9sm16	zea mays (maize). sbp-dc
sp_bacteria:Q9ZN36	+	7.00	100.28	489.86	425	Q9zn36	helicobacter pylori j99
sp_bacteria:Q9RYM5	+	7.00	100.28	489.23	428	Q9rym5	deinococcus radiodurans.
sp_invertebrate:Q9VYA9	-	7.00	100.24	489.23	428	!Q9vya9	drosophila melanogaste
sp_bacteria:Q9PBC1	+	7.00	100.21	488.81	430	Q9pbc1	xylella fastidiosa. thre
sp_bacteria:Q9KYS0	+	7.00	100.21	488.81	430	Q9kys0	streptomyces coelicolor.
sp_rodent:054844	+	7.00	100.15	487.98	434	054844	mus musculus (mouse). et
sp_virus:Q84402	+	7.00	100.13	487.77	435	Q84402	peanut bud necrosis viru
sp_fungi:Q9P452	+	7.00	100.13	487.77	435	Q9p452	sclerotinia sclerotiorum
sp_plant:Q9SM17	-	7.00	100.13	487.77	435	Q9sm17	zea mays (maize). sbp-dc
sp_plant:022778	+	7.00	100.12	487.56	436	022778	arabidopsis thaliana (mc
sp_bacteria:Q9X8S1	+	7.00	100.09	487.15	438	Q9x8s1	streptomyces coelicolor.
sp_bacteria:082244	+	7.00	100.08	486.95	439	082244	arabidopsis thaliana (mc
sp_human:Q9NP19	+	7.00	100.05	485.53	446	Q9np19	homo sapiens (human). ac
sp_plant:Q9LVA0	+	7.00	99.98	485.53	446	Q9lva0	arabidopsis thaliana (mc
sp_bacteria:P74824	+	7.00	99.96	485.33	447	P74824	sphingomonas sp. s88. hy
sp_bacteria:Q9X7Z6	+	7.00	99.96	485.33	447	Q9x7z6	streptomyces coelicolor.
sp_bacteria:Q9KXC8	+	7.00	99.93	484.94	449	Q9kxc8	escherichia coli O157:h7
sp_human:043934	+	7.00	99.93	484.94	449	043934	homo sapiens (human). ac
sp_human:Q9NXI5	+	7.00	99.93	484.74	450	Q9nxi5	homo sapiens (human). co
sp_bacteria:025596	+	7.00	99.92	484.74	450	025596	helicobacter pylori (cam
sp_bacteria:Q9L554	+	7.00	99.92	484.74	450	Q9l554	streptomyces galilaeus.
sp_human:Q9NP20	+	7.00	99.90	484.54	451	Q9np20	homo sapiens (human). ac
sp_plant:Q9M1X6	+	7.00	99.89	484.54	451	Q9m1x6	arabidopsis thaliana (mc
sp_plant:Q9M9D7	+	7.00	99.85	484.34	452	Q9m9d7	arabidopsis thaliana (mc
sp_bacteria:068903	+	7.00	99.85	483.75	455	068903	actinosynnema pretiosum
sp_human:Q9UP88	+	7.00	99.83	483.75	455	Q9up88	homo sapiens (human). ac
sp_bacteria:Q9KUF5	-	7.00	99.83	483.56	456	Q9kuf5	vibrio cholerae. proteas
sp_human:Q9Y215	+	7.00	99.83	483.56	456	Q9y215	homo sapiens (human). ac
sp_invertebrate:Q20084	-	7.00	99.83	483.56	456	!Q20084	caenorhabditis elegans
sp_archea:Q9YEU2	+	7.00	99.81	483.17	458	Q9yeu2	aeropyrum pernix. 458aa
sp_bacteria:066229	+	7.00	99.81	483.17	458	066229	escherichia coli. phosph
sp_bacteria:Q9RMU4	+	7.00	99.81	483.17	458	Q9rmu4	klebsiella pneumoniae. n
sp_human:043464	+	7.00	99.81	483.17	458	043464	homo sapiens (human). se
sp_rodent:Q9R108	+	7.00	99.81	483.17	458	Q9r108	mus musculus (mouse). se
sp_rodent:Q9JIY5	+	7.00	99.79	482.98	459	Q9jiy5	mus musculus (mouse). se
sp_bacteria:Q9ZCF1	+	7.00	99.79	482.98	459	Q9zcf1	rickettsia prowazekii.
sp_bacteria:Q9XII8	+	7.00	99.79	482.98	459	Q9xii8	streptomyces coelicolor.
sp_plant:Q9LGY8	+	7.00	99.78	482.78	460	Q9lgy8	oryza sativa (rice). sim
sp_bacteria:066236	+	7.00	99.78	482.78	460	066236	escherichia coli. phosph
sp_bacteria:066243	+	7.00	99.78	482.78	460	066243	klebsiella pneumoniae. p
sp_bacteria:Q59428	+	7.00	99.78	482.78	460	Q59428	escherichia coli. and kl
sp_organelle:Q9XKX4	-	7.00	99.78	482.78	460	Q9xkx4	chrysemys picta. nadh de
sp_organelle:Q9XPH7	-	7.00	99.78	482.78	460	Q9xph7	chelonia mydas caranigra
sp_virus:079790	+	7.00	99.78	482.78	460	079790	human immunodeficiency
sp_virus:Q69887	-	7.00	99.71	481.82	465	Q69887	human immunodeficiency v
sp_bacteria:069954	+	7.00	99.68	481.63	466	069954	streptomyces coelicolor.
sp_bacteria:Q53246	+	7.00	99.70	481.63	466	Q53246	rickettsia tsutsugamushi
sp_bacteria:Q53247	+	7.00	99.70	481.63	466	Q53247	rickettsia tsutsugamushi
sp_bacteria:Q53249	+	7.00	99.70	481.63	466	Q53249	rickettsia tsutsugamushi
sp_bacteria:Q53251	+	7.00	99.70	481.63	466	Q53251	rickettsia tsutsugamushi
sp_bacteria:005335	+	7.00	99.70	481.63	466	005335	rickettsia typhi. 47 kda
sp_invertebrate:076322	+	7.00	99.70	481.44	467	!076322	loligo pealeii (longfi
sp_bacteria:Q9L5J6	+	7.00	99.68	481.06	469	Q9l5j6	salmonella typhi. hypoth
sp_bacteria:Q9X803	-	7.00	99.66	481.06	469	Q9x803	oryza sativa (rice). sim
sp_bacteria:006604	-	7.00	99.66	481.06	469	006604	mycobacterium tuberculos
sp_plant:Q9S9A3	-	7.00	99.66	481.06	469	Q9s9a3	lycopersicon esculentum
sp_bacteria:Q9K7G0	+	7.00	99.58	479.94	475	Q9k7g0	bacillus halodurans. leu
sp_invertebrate:Q21555	-	7.00	99.58	479.94	475	!Q21555	caenorhabditis elegans

Right table

ID		7.00	%id	score	len	acc	description
sp_bacteria:066573	-	7.00	99.56	479.75	476	066573	aquilex aeolicus. ald
sp_bacteria:083837		7.00	99.52	479.20	479	083837	treponema pallidum. h
sp_human:Q92743		7.00	99.51	479.02	480	Q92743	homo sapiens (human).
sp_rodent:Q9R118		7.00	99.51	479.02	480	Q9R118	mus musculus (mouse).
sp_rodent:Q9QZK6		7.00	99.51	479.02	480	Q9QZK6	mus musculus (mouse)
sp_bacteria:Q9QZK5		7.00	99.50	478.83	480	Q9QZK5	ratus norvegicus (ra
sp_bacteria:066144		7.00	99.48	478.65	481	066144	citrobacter amalonati
sp_invertebrate:Q9RWF2		7.00	99.47	478.47	482	Q9rwf2	deinococcus radiodura
sp_invertebrate:051131					483	051131	borrelia burgdorferi
sp_invertebrate:P91239	+	7.00	99.44	478.10	485	!P91239	caenorhabditis cleg
sp_invertebrate:Q9V9N1	+	7.00	99.44	478.10	485	!Q9v9n1	drosophila melanoga
sp_bacteria:Q9RFC6		7.00	99.43	477.92	486	Q9rfc6	rhodobacter sphaeroid
sp_bacteria:050436		7.00	99.40	477.56	488	050436	mycobacterium tubercu
sp_mammal:Q28425		7.00	99.31	476.31	495	Q28425	gorilla gorilla goril
sp_plant:Q9SCN2		7.00	99.27	475.78	498	Q9scn2	arabidopsis thaliana
sp_invertebrate:076323	-	7.00	99.21	474.91	503	!076323	loligo pealii (lon
sp_virus:082010	+	7.00	99.21	474.91	503	082010	human papillomavirus
sp_bacteria:Q92C68		7.00	99.19	474.39	505	Q92c68	yersinia pestis. arud
sp_rodent:Q62741		7.00	99.17	474.39	506	Q62741	rattus norvegicus (ra
sp_bacteria:066973		7.00	99.12	473.53	510	066973	aquilex aeolicus. l-a
sp_virus:Q91DV8		7.00	99.11	473.53	511	Q91dv8	human immunodeficienc
sp_invertebrate:Q9U023	-	7.00	99.10	473.36	512	!Q9u023	giardia lamblia (gi
sp_bacteria:Q9KLS4		7.00	99.09	473.19	513	Q9kls4	vibrio cholerae. sod1
sp_fungi:Q97P3		7.00	99.09	473.19	513	Q97p3	schizosaccharomyces p
sp_bacteria:Q9RBP2		7.00	99.07	473.02	514	Q9rbp2	rhodococcus sp. ad45.
sp_plant:022509		7.00	99.05	472.69	516	022509	arabidopsis thaliana
sp_plant:Q9SYC2		7.00	99.05	472.69	516	Q9syc2	arabidopsis thaliana
sp_plant:Q9ZU50		7.00	99.04	472.52	517	Q9zu50	arabidopsis thaliana
sp_archea:027324		7.00	99.01	472.18	519	027324	methanobacterium ther
sp_bacteria:P72042		7.00	99.01	472.18	519	P72042	mycobacterium tubercu
sp_bacteria:Q9PCG3		7.00	99.01	472.18	519	Q9pcg3	xylella fastidiosa. 2
sp_virus:Q9YQ22		7.00	99.01	472.18	519	Q9yq22	bovine calicivirus. s
sp_plant:048881		7.00	98.96	471.51	523	048881	brassica napus (rape)
sp_plant:Q9LY45		7.00	98.91	470.85	527	Q9ly45	arabidopsis thaliana
sp_bacteria:069547		7.00	98.88	470.36	530	069547	mycobacterium leprae.
sp_bacteria:Q9R9Y7		7.00	98.88	470.36	530	Q9r9y7	mycobacterium leprae.
sp_plant:Q9SK21		7.00	98.87	470.20	531	Q9sk21	arabidopsis thaliana
sp_fungi:Q49972		7.00	98.84	469.87	534	Q49972	mycobacterium leprae.
sp_fungi:074996		7.00	98.83	469.87	534	074996	yarrowia lipolytica.
sp_invertebrate:Q9W0L7	+	7.00	98.82	469.55	535	Q9w0l7	drosophila melanoga
sp_bacteria:056650		7.00	98.81	469.39	536	056650	adeno-associated viru
sp_invertebrate:Q9VPM5	+	7.00	98.77	468.91	539	Q9vpm5	drosophila melanoga
sp_virus:066287		7.00			539	066287	unidentified. capsid
sp_virus:Q9YQS0		7.00	98.77	468.91	539	Q9yqs0	human calicivirus. ca
sp_virus:Q68537		7.00	98.77	468.91	539	Q68537	human calicivirus. an
sp_virus:Q9WRZ1		7.00	98.77	468.91	539	Q9wrz1	norwalk-like virus sp
sp_virus:Q9WRZ0		7.00	98.77	468.91	539	Q9wrz0	norwalk-like virus sp
sp_virus:Q9WRY9		7.00	98.77	468.91	539	Q9wry9	norwalk-like virus sp
sp_virus:Q9WRY8		7.00	98.77	468.91	539	Q9wry8	norwalk-like virus sp
sp_virus:Q9W9Y9		7.00	98.77	468.91	539	Q9w9y9	norwalk-like virus sp
sp_virus:Q9W9T8		7.00	98.77	468.91	539	Q9w9t8	norwalk-like virus sp
sp_virus:Q91V44		7.00	98.77	468.91	539	Q91v44	human calicivirus hu/
sp_virus:Q91V38		7.00	98.76	468.75	540	Q91v38	human calicivirus hu/
sp_bacteria:005692		7.00	98.75	468.75	540	005692	rhizobium leguminosar
sp_human:060230		7.00	98.76	468.75	540	060230	homo sapiens (human).
sp_human:Q14053		7.00	98.74	468.43	543	Q14053	homo sapiens (human).
sp_bacteria:Q52413		7.00	98.72	468.27	543	Q52413	pseudomonas aeruginos
sp_fungi:013449		7.00	98.49	465.16	563	013449	saccharomyces cerevis
sp_plant:023944		7.00	98.46	464.71	566	023944	fragaria vesca. cysta
sp_plant:Q40451		7.00	98.45	464.56	567	Q40451	nicotiana tabacum (cc
sp_invertebrate:Q18041	+	7.00	98.67	467.48	548	!Q18041	caenorhabditis eleg
sp_bacteria:066291		7.00	98.65	467.32	549	066291	mycobacterium tubercu
sp_bacteria:086752		7.00	98.61	466.70	553	086752	streptomyces coelicol
sp_invertebrate:Q24612	-	7.00	98.52	465.47	561	!Q24612	drosophila pseudoot
sp_plant:Q52413		7.00	98.49	465.16	563	Q52413	pseudomonas sp. alpha
sp_plant:Q9SV44		7.00	98.46	464.71	566	Q9sv44	arabidopsis thaliana
sp_plant:Q9R9H2		7.00	98.45	464.56	567	Q9r9h2	pseudomonas chlororap
sp_fungi:P78871		7.00	98.45	464.56	567	P78871	schizosaccharomyces f
sp_phage:Q9KJK0		7.00	98.44	464.41	568	Q9kjk0	bacteriophage 93w, a
sp_archea:028188		7.00	98.43	464.26	569	028188	archaeoglobus fulgidu
sp_bacteria:Q9KA71		7.00	98.40	463.96	571	Q9ka71	bacillus halodurans.
sp_plant:024022		7.00	98.40	463.96	571	024022	lycopersicon esculent

```
sp_invertebrate:Q9W2P2 +  7.00  98.39 463.81   ! Q9w2p2 drosophila melanogaste   572
sp_invertebrate:Q9N4Q1 +  7.00  98.37 463.51   ! Q9n4q1 caenorhabditis elegans   574
sp_bacteria:Q9K3P9     -  7.00  98.34 463.07     Q9k3p9 streptomyces coelicolor.  577
sp_bacteria:Q9X6M9     -  7.00  98.29 462.48     Q9x6m9 bartonella clarridgeiae.  581
sp_invertebrate:Q23318 +  7.00  98.29 462.48   ! Q23318 caenorhabditis elegans   581
sp_bacteria:Q9GR4      -  7.00  98.27 462.19     Q9gr4 bordetella bronchiseptic   583
sp_bacteria:Q9RGR3     -  7.00  98.27 462.19     Q9rgr3 bordetella bronchiseptic  583
sp_invertebrate:Q9X66  +  7.00  98.27 462.19   ! Q9x66 caenorhabditis elegans     583
sp_plant:P94017        +  7.00  98.26 462.04     P94017 arabidopsis thaliana (mc  584
sp_plant:Q9LMD2        -  7.00  98.25 461.90     Q9lmd2 arabidopsis thaliana (mc  585
sp_bacteria:Q9KLS8     -  7.00  98.23 461.61     Q9kls8 vibrio cholerae. methyl   587
sp_invertebrate:O00904 +  7.00  98.23 461.61   ! O00904 leishmania major. 1417    587
sp_plant:Q9SVE0        -  7.00  98.23 461.61     Q9sve0 arabidopsis thaliana (mc  587
sp_vertebrate:Q9YHZ0   -  7.00  98.22 461.46     Q9yhz0 gallus gallus (chicken).  588
sp_invertebrate:Q9XAH5 -  7.00  98.22 461.46     Q9xah5 streptomyces coelicolor.  588
sp_invertebrate:O02007 -  7.00  98.22 461.46   ! O02007 drosophila similans (f    588
sp_plant:Q92BR4        -  7.00  98.21 461.32     Q92br4 streptomyces coelicolor.  589
sp_plant:Q38801        -  7.00  98.20 461.17     Q38801 arabidopsis thaliana (mc  590
sp_rodent:Q90WJ6       -  7.00  98.20 461.17     Q90wj6 rattus norvegicus (rat)   590
sp_invertebrate:O77267 +  7.00  98.17 460.89   ! O77267 drosophila melanogaste    592
sp_archea:Q9UYA4       -  7.00  98.14 460.46     Q9uya4 pyrococcus abyssi. hypot  595
sp_bacteria:Q9PEF6     +  7.00  98.13 460.32     Q9pef6 xylella fastidiosa. succ  596
sp_rodent:O35727       -  7.00  98.12 460.17     O35727 mus musculus (mouse). fa  597
sp_human:O75017        +  7.00  98.11 460.03     O75017 homo sapiens (human). le  598
sp_plant:Q9SA65        +  7.00  98.10 459.89     Q9sa65 arabidopsis thaliana (mc  599
sp_fungi:Q9P8M8        -  7.00  98.04 459.05     Q9p8m8 aspergillus niger. gluco  605
sp_invertebrate:Q9TVK2 +  7.00  97.99 458.49   ! Q9tvk2 caenorhabditis elegans    609
sp_invertebrate:Q9YM61 +  7.00  97.97 458.22   ! Q9ym61 drosophila melanogaste    611
sp_plant:Q9SFX3        -  7.00  97.94 457.80     Q9sfx3 arabidopsis thaliana (mc  614
sp_vertebrate:Q9W721   -  7.00  97.92 457.53     Q9w721 brachydanio rerio (zebra  616
sp_bacteria:Q9S6T5     -  7.00  97.91 457.39     Q9s6t5 streptomyces coelicolor.  617
sp_virus:O56651        -  7.00  97.87 456.85     O56651 adeno-associated virus 2  621
sp_virus:O89268        -  7.00  97.87 456.85     O89268 adeno-associated virus 2  621
sp_fungi:Q06451        -  7.00  97.86 456.71     Q06451 saccharomyces cerevisiae  622
sp_invertebrate:Q9W4A2 -  7.00  97.85 456.58   ! Q9w4a2 drosophila melanogaste    623
sp_virus:O56136        -  7.00  97.85 456.58     O56136 adeno-associated virus 6  623
sp_virus:Q9WBP7        -  7.00  97.85 456.58     Q9wbp7 adeno-associated virus 1  623
sp_bacteria:Q9ZUA3     +  7.00  97.79 455.77     Q9p5m2 neurospora crassa. conse  629
sp_fungi:Q9P5M2        +  7.00  97.73 454.98     Q9p5m2 neurospora crassa. conse  635
sp_mammal:Q28423       +  7.00  97.73 454.98     Q28423 gorilla gorilla gorilla   635
sp_human:Q9NSG0        -  7.00  97.65 453.93     Q9nsg0 homo sapiens (human). hy  643
sp_bacteria:Q9KKB9     -  7.00  97.63 453.67     Q9kkb9 escherichia coli o157:h7  645
sp_invertebrate:Q9NEE6 -  7.00  97.63 453.67   ! Q9nee6 leishmania major. hypo    645
sp_phage:Q9XJK8        +  7.00  97.63 453.67     Q9xjk8 bacteriophage 933w, and   645
sp_bacteria:P70997     -  7.00  97.61 453.41     P70997 bacillus subtilis. hypot  647
sp_bacteria:Q9KKQ7     -  7.00  97.59 453.16     Q9kkq7 vibrio cholerae. pts sys  649
sp_invertebrate:Q20442 +  7.00  97.59 453.16   ! Q20442 caenorhabditis elegans    649
sp_invertebrate:O16242 +  7.00  97.57 452.90   ! O16242 caenorhabditis elegans    651
sp_bacteria:Q928C4     +  7.00  97.55 452.64     Q928c4 chlamydia pneumoniae (CH  653
sp_bacteria:Q9SF0      -  7.00  97.55 452.64     Q9sf0 chlamydia pneumoniae (CH   653
sp_virus:O93130        +  7.00  97.47 451.63     O93130 cowpox virus (cpv). d81   661
sp_plant:Q9M2R4        +  7.00  97.46 451.51     Q9m2r4 arabidopsis thaliana (mc  662
sp_invertebrate:P81538 +  7.00  97.38 450.39   ! P81538 drosophila ananassae (    671
sp_invertebrate:Q9N9L7 +  7.00  97.38 450.39   ! Q9n9l7 leishmania major. hypo    671
sp_fungi:Q9Y7A7        -  7.00  97.29 449.28     Q9y7a7 exophiala dermatitidis (  680
sp_invertebrate:Q9W531 +  7.00  97.14 447.37     Q9w531 bacillus haloourans. cad  696
sp_bacteria:O34003     +  7.00  97.11 446.89     O34003 rhodobacter sphaeroides   700
sp_invertebrate:Q9VBZ5 +  7.00  97.11 446.89     Q9vbz5 drosophila melanogaste    700
sp_organelle:O9TLE9    +  7.00  97.08 446.54     Q9tle9 nemesia strumosa. nadh    703
sp_bacteria:Q9RZF3     +  7.00  97.08 446.54     Q9rzf3 deinococcus radiodurans   703
sp_invertebrate:O16241 +  7.00  97.08 446.54   ! O16241 caenorhabditis elegans    704
sp_rodent:Q9WVM9       -  7.00  97.07 446.43     Q9wvm9 mus musculus (mouse). ly  707
sp_invertebrate:Q9N9U0 +  7.00  97.04 446.03   ! Q9n9u0 caenorhabditis elegans    710
sp_bacteria:Q9K5Q2     +  7.00  97.02 445.73     Q9k5q2 bacillus haloourans. cac  712
sp_bacteria:O85906     -  7.00  97.00 445.50     O85906 sphingomonas aromaticiv   722
sp_rodent:O54729       -  7.00  96.91 444.36     O54729 rattus norvegicus (rat).  726
sp_invertebrate:Q9U0Z4 +  7.00  96.88 443.91   ! Q9u0z4 leishmania major. hypo    726
sp_plant:Q9LMK3        -  7.00  96.87 443.79     Q9lmk3 arabidopsis thaliana (mc  727
sp_vertebrate:Q9W7F9   -  7.00  96.85 443.51     Q9w7f9 gallus gallus (chicken).  729
sp_phage:Q9ZXA5        -  7.00  96.81 443.01     Q9zxa5 bacteriophage phi-c31.    734
sp_bacteria:Q9KNQ3     -  7.00  96.77 442.57     Q9knq3 vibrio cholerae. primoso  738
sp_rodent:Q9QY02       -  7.00  96.77 442.57     Q9qy02 rattus norvegicus (rat)   738
```

```
 739  442.46  96.76  7.00  +  sp_archea:Q9UWN2        ! Q9uwn2 thermococcus sp. b10
 739  442.46  96.76  7.00  -  sp_plant:O65420           O65420 arabidopsis thaliana
 744  441.91  96.72  7.00  -  sp_plant:Q9MB58           Q9mb58 arabidopsis thaliana
 745  441.80  96.71  7.00  -  sp_plant:Q9SP17           Q9sp17 arabidopsis thaliana
 745  441.80  96.71  7.00  -  sp_plant:Q9LL39           Q91139 arabidopsis thaliana
 749  441.37  96.68  7.00  -  sp_bacteria:P73687        P73687 synechocystis sp. (s
 752  441.04  96.65  7.00  -  sp_bacteria:P96316        P96316 cellvibrio gilvus. b
 753  440.93  96.64  7.00  -  sp_human:O94993           O94993 homo sapiens (human)
 754  440.83  96.64  7.00  +  sp_plant:Q9LFF1           Q9lff1 arabidopsis thaliana
 769  439.23  96.51  7.00  +  sp_fungi:O74253           O74253 pycnoporus cinnabari
 769  439.23  96.51  7.00  +  sp_invertebrate:Q20440  ! Q20440 caenorhabditis ele
 782  437.88  96.36  7.00  +  sp_human:Q9NVB2           Q9nvb2 homo sapiens (human)
 787  437.37  96.36  7.00  +  sp_virus:Q9WFB5           Q9wfb5 snow goose hepatitis
 787  437.37  96.36  7.00  +  sp_virus:Q9WFB2           Q9wfb2 snow goose hepatitis
 787  437.37  96.36  7.00  +  sp_virus:Q9WFA8           Q9wfa8 snow goose hepatitis
 787  437.37  96.36  7.00  +  sp_virus:Q9WFA5           Q9wfa5 snow goose hepatitis
 787  437.37  96.36  7.00  +  sp_virus:Q9WFA2           Q9wfa2 snow goose hepatitis
 793  436.75  96.31  7.00  +  sp_invertebrate:Q9VW27  ! Q9vw27 drosophila melanog
 793  436.75  96.31  7.00  +  sp_plant:Q9LIT9           Q91it9 oryza sativa (rice).
 796  436.45  96.29  7.00  -  sp_invertebrate:O16145  ! O16145 caenorhabditis ele
 799  436.15  96.27  7.00  -  sp_plant:O23190           O23190 arabidopsis thaliana
 804  435.65  96.23  7.00  -  sp_invertebrate:Q9VAK6  ! Q9vak6 drosophila melanog
 807  435.35  96.20  7.00  -  sp_plant:Q9XI07           Q9xi07 arabidopsis thaliana
 809  435.15  96.19  7.00  -  sp_invertebrate:O44969  ! O44969 caenorhabditis ele
 813  434.76  96.16  7.00  -  sp_bacteria:O55799        O55799 synechocystis sp. (s
 838  432.34  95.96  7.00  -  sp_invertebrate:P90882  ! P90882 caenorhabditis ele
 839  432.25  95.96  7.00  -  sp_invertebrate:Q10652  ! Q10652 caenorhabditis ele
 841  432.06  95.94  7.00  +  sp_plant:Q9SRR2           Q9srr2 arabidopsis thaliana
 843  431.87  95.93  7.00  -  sp_invertebrate:O61123  ! O61123 entamoeba histolyt
 846  431.59  95.90  7.00  -  sp_plant:O49507           O49507 arabidopsis thaliana
 847  431.49  95.90  7.00  -  sp_plant:P93068           P93068 brassica oleracea (c
 847  431.49  95.90  7.00  -  sp_plant:O04383           O04383 brassica oleracea (c
 852  431.03  95.86  7.00  +  sp_invertebrate:Q9U4G5  ! Q9u4g5 drosophila melanog
 854  430.84  95.84  7.00  +  sp_fungi:Q9Y6Z8           Q9y6z8 aspergillus oryzae.
 855  430.75  95.84  7.00  +  sp_fungi:Q9RDL9           Q9rdl9 streptomyces coelico
 860  430.29  95.80  7.00  -  sp_fungi:Q02290           Q02290 neocallimastix patri
 874  429.01  95.70  7.00  +  sp_invertebrate:O18109  ! O18109 caenorhabditis ele
 878  428.65  95.67  7.00  +  sp_invertebrate:P91859  ! P91859 caenorhabditis rou
 878  428.65  95.67  7.00  -  sp_virus:Q83101           Q83101 unidentified leishma
 879  428.56  95.67  7.00  -  sp_virus:Q83101           Q83101 unidentified leishman
 881  428.38  95.64  7.00  -  sp_rodent:Q9WV91          Q9wv91 mus musculus (mouse)
 885  428.02  95.62  7.00  -  sp_fungi:O13981           O13981 schizosaccharomyces
 888  427.76  95.59  7.00  +  sp_plant:Q9ZQX6           Q9zqx6 arabidopsis thaliana
 893  427.31  95.55  7.00  -  sp_bacteria:O53121        O53121 mycobacterium leprae
 897  426.96  95.53  7.00  -  sp_human:Q9NRY3           Q9nry3 homo sapiens (human)
 900  426.70  95.51  7.00  -  sp_fungi:Q9P4D3           Q9p4d3 zygosaccharomyces rou
 905  426.27  95.47  7.00  -  sp_invertebrate:O18260  ! O18260 caenorhabditis eleg
 915  425.41  95.40  7.00  -  sp_bacteria:O32863        O32863 mycoplasma hominis. F
 916  425.32  95.40  7.00  -  sp_invertebrate:Q9ROX2    Q9rux2 deinococcus radiodura
 921  424.89  95.36  7.00  -  sp_plant:Q9LMP6           Q91mp6 arabidopsis thaliana
 922  424.81  95.36  7.00  +  sp_invertebrate:O46354  ! O46354 caenorhabditis eleg
 924  424.64  95.34  7.00  -  sp_human:Q9P2B2           Q9p2b2 homo sapiens (human)
 942  423.14  95.22  7.00  -  sp_vertebrate:O57452      O57452 fundulus heteroclitus
 944  422.97  95.21  7.00  -  sp_invertebrate:Q9VCE7  ! Q9vce7 drosophila melanoga
 946  422.81  95.19  7.00  -  sp_virus:Q69139           Q69139 human herpesvirus 4 (
 955  422.07  95.13  7.00  -  sp_plant:Q22913           Q22913 arabidopsis thaliana
 957  421.91  95.12  7.00  -  sp_human:O60282           O60282 homo sapiens (human)
 968  421.02  95.05  7.00  -  sp_invertebrate:O18359  ! O18359 drosophila melanoga
 973  420.62  95.07  7.00  -  sp_invertebrate:Q10953  ! Q10953 caenorhabditis eleg
 980  420.07  94.97  7.00  +  sp_human:O75179           O75179 homo sapiens (human)
 990  419.28  94.90  7.00  -  sp_virus:O76351         ! O76351 giardia lamblia (gi
 995  418.90  94.87  7.00  -  sp_invertebrate:Q9W2L7    Q9w2l7 drosophila melanoga
1003  418.28  94.82  7.00  -  sp_bacteria:P95542        P95542 pseudomonas putida.
1007  417.97  94.79  7.00  -  sp_human:O43424           O43424 homo sapiens (human)
1008  417.90  94.79  7.00  -  sp_rodent:O61625          Q61625 mus musculus (mouse)
1033  416.02  94.63  7.00  -  sp_rodent:O63226          O63226 rattus norvegicus (r
1035  415.87  94.62  7.00  -  sp_plant:Q9XEK5           Q9xek5 chlamydomonas reinha
1035  415.87  94.62  7.00  -  sp_rodent:O35911          O35911 mus musculus (mouse)
1036  415.80  94.61  7.00  -  sp_bacteria:O08684        O08684 cricetulus griseus (
1036  415.80  94.61  7.00  -  sp_rodent:O35856          O35856 rattus norvegicus (r
```

ID		score		rank	description
sp_rodent:O08959	7.00	94.61	415.80	1036	O08959 rattus norvegicus (rat)
sp_invertebrate:Q9VZ30 +	7.00	94.59	415.50	1040	Q9vz30 drosophila melanogast
sp_rodent:Q90840 -	7.00	94.58	415.35	1040	Q90840 gallus gallus (chicken)
sp_plant:Q9LM01	7.00	94.54	414.92	1042	Q9lm01 zea mays (maize). calci
sp_invertebrate:O45962 +	7.00	94.46	413.90	1048	O45962 caenorhabditis elegan
sp_invertebrate:Q9VIV2 +	7.00	94.41	413.40	1062	Q9viv2 drosophila melanogast
sp_bacteria:Q9JY62	7.00	94.41	413.40	1069	Q9jy62 neisseria meningitidis
sp_human:Q9UF98	7.00	94.41	413.40	1069	Q9uf98 homo sapiens (human). h
sp_rodent:O54765 -	7.00	94.38	413.05	1074	O54765 rattus norvegicus (rat)
sp_rodent:Q92280	7.00	94.38	413.05	1074	Q92280 mus musculus (mouse). F
sp_invertebrate:Q19192 -	7.00	94.37	412.84	1077	Q19192 caenorhabditis elegan
sp_invertebrate:Q9VAM3 +	7.00	94.33	412.35	1084	Q9vam3 drosophila melanogast
sp_human:O75808	7.00	94.31	412.21	1086	O75808 homo sapiens (human). s
sp_plant:Q9SM80	7.00	94.26	411.58	1089	Q9sm80 oryza sativa (rice). zh
sp_rodent:O35867	7.00	94.26	411.58	1095	O35867 rattus norvegicus (rat)
sp_invertebrate:Q9LG8	7.00	94.26	411.58	1095	Q9jlg8 mus musculus (mouse). s
sp_bacteria:O51934	7.00	94.21	410.96	1104	O51934 thermotoga maritima. re
sp_invertebrate:Q9NED2 +	7.00	94.15	410.28	1114	Q9ned2 leishmania major. hyp
sp_human:Q9NM5	7.00	94.14	410.14	1116	Q9nqm5 homo sapiens (human). h
sp_fungi:Q12734	7.00	94.11	409.81	1121	Q12734 saccharomyces cerevisia
sp_invertebrate:Q9U2C0 +	7.00	94.09	409.61	1124	Q9u2c0 caenorhabditis elegan
sp_fungi:Q9P664	7.00	94.05	409.07	1132	Q9p664 neurospora crassa. rela
sp_fungi:Q12139	7.00	94.04	409.01	1133	Q12139 saccharomyces cerevisia
sp_human:Q13082	7.00	93.94	407.76	1152	Q13082 homo sapiens (human). n
sp_vertebrate:Q90783 +	7.00	93.82	406.40	1173	Q90783 gallus gallus (chicken)
sp_plant:O65507	7.00	93.79	406.02	1179	O65507 arabidopsis thaliana (m
sp_invertebrate:Q94447 +	7.00	93.77	405.77	1183	Q94447 calliphora vicina (bl
sp_human:Q90NG7	7.00	93.64	404.27	1207	Q9u0ng7 rattus norvegicus (rat)
sp_rodent:Q9QX70 +	7.00	93.63	404.15	1209	Q9qx70 rattus norvegicus (rat)
sp_human:O60305	7.00	93.62	403.96	1212	O60305 homo sapiens (human). k
sp_invertebrate:Q17240 +	7.00	93.59	403.66	1217	Q17240 bombyx mori (silk mot
sp_invertebrate:Q18270 +	7.00	93.58	403.60	1218	Q18270 caenorhabditis elegan
sp_invertebrate:Q77285 -	7.00	93.50	402.63	1234	Q77285 drosophila melanogast
sp_invertebrate:Q9W0B8	7.00	93.50	402.63	1234	Q9w0b8 drosophila melanogast
sp_invertebrate:O69947	7.00	93.48	402.39	1238	O69947 streptomyces coelicolor
sp_invertebrate:Q10908 -	7.00	93.30	401.21	1258	Q10908 caenorhabditis elegan
sp_invertebrate:Q9VI7I7 -	7.00	93.28	400.10	1277	Q9vi7i7 drosophila melanogast
sp_plant:Q64496	7.00	92.95	396.30	1345	Q64496 arabidopsis thaliana (m
sp_invertebrate:Q9NEB8 +	7.00	92.91	395.87	1353	Q9neb8 gallus gallus (chicken)
sp_virus:P89442	7.00	92.90	395.71	1356	P89442 herpes simplex virus (t
sp_invertebrate:Q23371 -	7.00	92.82	394.75	1374	Q23371 caenorhabditis elegan
sp_invertebrate:Q9W3A4 +	7.00	92.62	392.52	1410	Q9w3a4 drosophila melanogast
sp_bacteria:O85862	7.00	92.59	391.71	1417	O85862 sphingomonas aromaticiv
sp_bacteria:Q9KA72	7.00	92.55	391.35	1425	Q9ka72 bacillus halodurans. dr
sp_plant:Q9SJA0	7.00	92.52	390.91	1433	Q9sja0 arabidopsis thaliana (m
sp_virus:Q9IDV9 +	7.00	92.48	390.34	1440	Q9idv9 human immunodeficiency
sp_human:Q9UQP1	7.00	92.34	389.34	1449	Q9uqp1 homo sapiens (human). h
sp_invertebrate:Q9V4C8 +	7.00	92.26	388.47	1481	Q9v4c8 drosophila melanogast
sp_invertebrate:Q21442 -	7.00	92.18	387.57	1499	Q21442 caenorhabditis elegan
sp_human:O94770	7.00	92.13	385.86	1518	O94770 homo sapiens (human). a
sp_human:Q9UEH6	7.00	91.85	383.84	1555	Q9ueh6 homo sapiens (human). h
sp_invertebrate:Q9VYW1 +	7.00	91.66	381.80	1600	Q9vyw1 drosophila melanogast
sp_human:Q12823	7.00	91.50	380.00	1647	Q12823 homo sapiens (human). a
sp_human:Q90MG6	7.00	91.50	380.00	1690	Q9omg6 homo sapiens (human). c
sp_human:Q9Y4L4	7.00	91.50	379.96	1690	Q9y4l4 homo sapiens (human). c
sp_plant:Q9SR02	7.00	91.43	379.26	1691	Q9sr02 arabidopsis thaliana (
sp_human:O94895	7.00	91.41	378.97	1703	O94895 homo sapiens (human). k
sp_human:Q9UJP5 +	7.00	91.39	378.81	1708	Q9ujp5 homo sapiens (human). c
sp_human:Q13029	7.00	91.34	378.20	1715	Q13029 mus musculus (mouse). m
sp_rodent:Q60592 -	7.00	91.30	377.76	1719	Q60592 mus musculus (mouse). m
sp_human:Q14675	7.00	91.19	376.66	1734	Q14675 homo sapiens (human). k
sp_human:Q9Y6B8	7.00	90.99	374.40	1745	Q9y6b8 homo sapiens (human). s
sp_bacteria:Q9L8C8	7.00	90.99	374.40	1773	Q9l8c8 sorangium cellulosum. e
sp_bacteria:Q9KIZ8	7.00	90.95	373.95	1832	Q9kiz8 sorangium cellulosum.
sp_invertebrate:Q9V9S7 +	7.00	90.94	373.87	1844	Q9v9s7 drosophila melanogast
sp_invertebrate:Q9VWN8 +	7.00	90.72	371.50	1846	Q9vwn8 drosophila melanogast
sp_invertebrate:Q9W4M7 -	7.00	90.68	371.11	1911	Q9w4m7 drosophila melanogast
sp_human:O43355	7.00	90.32	367.23	1922	O43355 homo sapiens (human). F
sp_human:Q92508	7.00	90.26	366.57	2035	Q92508 homo sapiens (human). d
sp_human:O75055	7.00	90.26	366.57	2055	O75055 homo sapiens (human). k
sp_plant:Q9LGG0	7.00	90.25	366.47	2058	Q9lgg0 oryza sativa (rice)
sp_invertebrate:Q9VWE6 +	7.00	90.16	365.53	2087	Q9vwe6 drosophila melanog
sp_invertebrate:O49624	7.00	90.06	364.54	2118	O49624 mycobacterium leprae
sp_rodent:Q92305	7.00	89.91	362.95	2169	Q92305 cavia porcellus (gui
sp_human:Q9UHP0	7.00	89.91	362.88	2171	Q9uhp0 homo sapiens (human)
sp_invertebrate:Q9NGQ2 +	7.00	89.81	361.85	2205	Q9ngq2 dictyostelium disc
sp_invertebrate:Q09515 +	7.00	89.80	361.76	2208	Q09515 caenorhabditis ele
sp_human:Q93072	7.00	89.79	361.64	2212	Q93072 homo sapiens (human)
sp_human:Q9Y5T3	7.00	89.70	360.71	2243	Q9y5t3 homo sapiens (human)
sp_human:Q9Y5T2	7.00	89.68	360.51	2250	Q9y5t2 homo sapiens (human)
sp_human:Q9NVU8	7.00	89.65	360.18	2261	Q9nvu8 homo sapiens (human)
sp_human:Q9NYU4	7.00	89.63	360.04	2266	Q9nyu4 homo sapiens (human)
sp_human:Q9NYV1	7.00	89.61	359.83	2273	Q9nyv1 homo sapiens (human)
sp_human:Q9NYU6	7.00	89.58	359.51	2284	Q9nyu6 homo sapiens (human)
sp_human:Q9NYU9	7.00	89.54	359.51	2298	Q9nyu9 homo sapiens (human)
sp_invertebrate:Q9N6Q1 -	7.00	89.54	359.02	2301	Q9n6q1 leishmania major.
sp_human:Q9NYU7	7.00	89.52	358.88	2306	Q9nyu7 homo sapiens (human)
sp_human:Q9NYV0	7.00	89.48	358.45	2321	Q9nyv0 homo sapiens (human)
sp_invertebrate:O16302 -	7.00	89.46	358.28	2327	O16302 caenorhabditis ele
sp_human:Q9NYU5	7.00	89.45	358.14	2332	Q9nyu5 homo sapiens (human)
sp_human:Q9UHN9	7.00	89.33	356.88	2377	Q9uhn9 homo sapiens (human)
sp_human:Q9NS88	7.00	88.99	353.43	2506	Q9ns88 homo sapiens (human)
sp_invertebrate:Q21980 +	7.00	88.93	352.78	2531	Q21980 caenorhabditis ele
sp_human:Q9NY15	7.00	88.83	351.79	2570	Q9ny15 homo sapiens (human)
sp_plant:Q9XIR5	7.00	88.52	348.66	2698	Q9xir5 arabidopsis thaliana
sp_invertebrate:Q9XZB8 -	7.00	88.50	348.38	2710	Q9xzb8 plasmodium falcipa
sp_invertebrate:Q9NKP7 +	7.00	87.99	343.28	2936	Q9nkp7 leishmania major.
sp_rodent:Q90XJ5 -	7.00	87.84	341.80	3006	Q9qxj5 mus musculus (mouse)
sp_bacteria:Q55582	7.00	87.79	341.32	3029	Q55582 synechocystis sp. (s
sp_virus:P89459	7.00	87.59	339.42	3122	P89459 herpes simplex virus
sp_human:Q9DQH3	7.00	86.70	330.73	3595	Q9uqh3 homo sapiens (human)
sp_vertebrate:Q9W6V0 -	7.00	86.66	330.37	3616	Q9w6v0 gallus gallus (chick
sp_invertebrate:P91904 -	7.00	86.51	328.91	3704	P91904 caenorhabditis eleg
sp_human:Q9Y6Y2	7.00	86.18	325.82	3899	Q9y6y2 homo sapiens (human)
sp_human:Q9UQQ4	7.00	86.17	325.69	3908	Q9uqq4 homo sapiens (human)
sp_human:Q99996	7.00	86.16	325.64	3911	Q99996 homo sapiens (human)
sp_rodent:Q61291	7.00	85.21	316.76	4545	Q61291 mus musculus (mouse)
sp_rodent:Q9QX6 +	7.00	84.81	313.20	4833	Q9qyx6 mus musculus (mouse)
sp_rodent:Q9JLT1 +	7.00	84.75	312.65	4880	Q9jlt1 rattus norvegicus (r
sp_rodent:Q9QX7 +	7.00	84.55	310.82	5038	Q9qyx7 mus musculus (mouse)
sp_bacteria:O52545 -	7.00	84.52	310.57	5060	O52545 amycolatopsis medite
sp_bacteria:O52789 -	7.00	84.51	310.47	5069	O52789 amycolatopsis medite
sp_rodent:Q9JKS6 -	7.00	84.49	310.29	5085	Q9jks6 rattus norvegicus (r
sp_invertebrate:Q9NI64 +	7.00	84.20	307.70	5322	Q9ni64 drosophila melanog
sp_bacteria:Q9R9J0 +	7.00	84.15	307.10	5369	Q9r9j0 bacillus subtilis."
sp_bacteria:O85168 +	7.00	80.60	275.72	9376	O85168 pseudomonas syringae
sp_bacteria:Q9L4X3 -	7.00	80.53	275.16	9477	Q9l4x3 streptomyces noursei

seq_name: sp_bacteria:Q9JVT1

seq_documentation_block:
```
ID   Q9JVT1       PRELIMINARY;       PRT;   499 AA.
AC   Q9JVT1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE PERIPLASMIC SERINE PROTEASE (EC 3.4.21.).
GN   NMA0710.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65599;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis 22491.";
```

```
RL   Nature 404:502-506(2000).
DR   EMBL; AL162754; CAB83996.1; -.
KW   Protease; Hydrolase.
SQ   SEQUENCE   499 AA;  52563 MW;  73E979F0A7F5B969 CRC64;

alignment_scores:
       Quality:   127.00                    Length:      127
         Ratio:    1.000                      Gaps:        0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Q9JVT1      ..

Align seg 1/1   to: Q9JVT1   from: 1 to: 499

583 TTGAAACCGGGCGAATGGTCGCTGCCATCGGCCGCCCTTCGGCTTGA     632
    |||||||||||||||||||||||||||||||||||||||||||||||
195 LeuLysProGlyGluTrpValAlaAlaIleGlyAlaProPheGlyPheAs  211

633 CAACAGCGGTGACCGCCGGCATCGTGTCCGCCAAAGGCAGAAGCCTGCCCA  682
    ||||||||||||||||||||||||||||||||||||||||||||||||||
211 pAsnSerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProA   228

683 ACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCCATCAATCCGGGC   732
    |||||||||||||||||||||||||||||||||||||||||||||||||
228 snGluSerTyrThrProPheIleGlnThrAspValAlaIleAsnProGly   244

733 AATTCCGGCCGGCCGCCGTGTTCAACTTAAAAGGACAGTCGTCGGCATCAA  782
    ||||||||||||||||||||||||||||||||||||||||||||||||||
245 AsnSerGlyGlyProLeuPheAsnLeuLysGlyValGlnValValGlyIleAs 261

783 TTCGCAAATATACAGCCGCCAGCGGCGGATTCATGGGCATCTCCTTTGCCA  832
    ||||||||||||||||||||||||||||||||||||||||||||||||||
261 nSerGlnIleTyrSerArgSerGlyGlyGlyPheMetGlyIleSerPheAlaI 278

833 TCCCGATTGACGTTGCCATGAATGTCGCCGAACAGCTGAAAAACACCGGC   882
    |||||||||||||||||||||||||||||||||||||||||||||||||
278 leProIleAspValAlaMetAsnValAlaGluGlnLeuLysAsnThrGly   294

883 AAAGTCCAACCGCGACAACTGGGCGTGATTATTCAGGAAGTATCCTACGG   932
    |||||||||||||||||||||||||||||||||||||||||||||||||
295 LysValGlnArgGlyGlnLeuGlyValIleIleGlnGluValSerTyrGl   311

933 TTTGGACAGTCGTTCGGTCTGGATAAAGCC  963
    ||||||||||||||||||||||||||||||
311 yLeuAlaGlnSerPheGlyLeuAspLysAla 321

seq_name: sp_bacteria:Q44476

seq_documentation_block:
ID   Q44476    PRELIMINARY;    PRT;    473 AA.
AC   Q44476;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   MUCD.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UW 136;
RX   MEDLINE=96178940; PubMed=8606151;
RA   Martinez-Salazar J.M., Moreno S., Najera R., Boucher J.C., Espin G.,
RA   Soberon-Chavez G., Deretic V.;
RT   "Characterization of the genes coding for the putative sigma factor
RT   AlgU and its regulators MucA, MucB, MucC, and MucD in Azotobacter
RT   vinelandii and evaluation of their roles in alginate biosynthesis.";
RL   J. Bacteriol. 178:1800-1808(1996).
DR   EMBL; U30799; AAB01513.1; -.
DR   INTERPRO; IPR001254;
```

```
DR   INTERPRO; IPR001478; -.
DR   PFAM; PF00089; trypsin; 2.
DR   PFAM; PF00595; PDZ; 2.
SQ   SEQUENCE   473 AA;  50282 MW;  C97B357D897738AB CRC64;

alignment_scores:
       Quality:    19.00                    Length:       19
         Ratio:    1.000                      Gaps:        0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Q44476      ..

Align seg 1/1   to: Q44476   from: 1 to: 473

697 CCCTTCATCCAAACCGACGTTGCCATCAATCCGGGCAATTCCGGCGGCCC   746
    ||||||||||||||||||||||||||||||||||||||||||||||||||
204 ProPheIleGlnThrAspValAlaIleAsnProGlyAsnSerGlyGlyPr   220

747 GCTGTTC   753
    |||||||
220 oLeuPhe   222

seq_name: sp_bacteria:Q57155

seq_documentation_block:
ID   Q57155    PRELIMINARY;    PRT;    474 AA.
AC   Q57155;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   MUCD.
GN   MUCD.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=93391358; PubMed=8378309;
RA   Martin D.W., Schurr M.J., Mudd M.H., Govan J.R., Holloway B.W.,
RA   Deretic V.;
RT   "Mechanism of conversion to mucoidy in Pseudomonas aeruginosa
RT   infecting cystic fibrosis patients.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:8377-8381(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=95286510; PubMed=7768826;
RA   Yu H., Schurr M.J., Deretic V.;
RT   "Functional equivalence of Escherichia coli sigma E and Pseudomonas
RT   aeruginosa AlgU: E. coli rpoE restores mucoidy and reduces sensitivity
RT   to reactive oxygen intermediates in algU mutants of P. aeruginosa.";
RL   J. Bacteriol. 177:3259-3268(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=96134987; PubMed=8550474;
RA   Boucher J.C., Martinez-Salazar J., Schurr M.J., Mudd M.H., Yu H.,
RA   Deretic V.;
RT   "Two distinct loci affecting conversion to mucoidy in Pseudomonas
RT   aeruginosa in cystic fibrosis encode homologs of the serine protease
RT   HtrA.";
RL   J. Bacteriol. 178:511-523(1996).
DR   EMBL; U49151; AAC43718.1; -.
DR   EMBL; U32853; AAC43676.1; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 2.
```

```
DR   PRINTS; PR00834; PROTEASES2C.
SQ   SEQUENCE   474 AA;  50321 MW;  8AB5D2A89867BEBE CRC64;

alignment_scores:
         Quality:   19.00          Length:      19
         Ratio:     1.000          Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Q57155

Align seg 1/1  to: Q57155  from: 1  to: 474

   637 AGGTGACCGCCGGCATCGTGTCCGCCCAAAGGCAGAAGCCTGCCCAACGA  686
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   184 SerValThrAlaGlyIleValSerAlaLysGlyArgSerLeuProAsnGl  200

   687 AAGCTAC  693
       |||||||
   200 uSerTyr  202

seq_name: sp_bacteria:Q9KJN6

seq_documentation_block:
ID   Q9KJN6       PRELIMINARY;   PRT;  500 AA.
AC   Q9KJN6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE SERINE PROTEASE DO-LIKE PRECURSOR.
GN   HTRA.
OS   Myxococcus xanthus.
OC   Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC   Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX   NCBI_TaxID=34;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DZ2;
RX   MEDLINE=20032045; PubMed=10564511;
RA   Cho K., Zusman D.R.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF163841; AAF87931.1; -.
KW   Protease.
SQ   SEQUENCE   500 AA;  50498 MW;  012DCAE1F38B9CDD CRC64;

alignment_scores:
         Quality:   13.00          Length:      13
         Ratio:     1.000          Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x Q9KJN6

Align seg 1/1  to: Q9KJN6  from: 1  to: 500

   718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCCGCTGTTCAAC  756
       ||||||||||||||||||||||||||||||||||||||||||
   245 AlaIleAsnProGlyAsnSerGlyGlyProLeuPheAsn  257

seq_name: sp_bacteria:035021

seq_documentation_block:
ID   035021       PRELIMINARY;   PRT;  224 AA.
AC   035021;
```

```
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   YVTB PROTEIN.
GN   YVTB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita Y., Fujita M., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98015415; PubMed=9353931;
RA   Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
RT   "Sequencing of regions downstream of addA (98 degrees) and citG (289
RT   degrees) in Bacillus subtilis.";
RL   Microbiology 143:3305-3308(1997).
DR   EMBL; Z99120; CAB15289.1; -.
DR   EMBL; Z93941; CAB07969.1; -.
DR   HSSP; P09331; 1AGJ.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 1.
SQ   SEQUENCE   224 AA;  23699 MW;  34A7CF0D79C7FB59 CRC64;

alignment_scores:
         Quality:   11.00          Length:      11
         Ratio:     1.000          Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x 035021
```

```
RN   [2]
RT   "Sporulation timing in Myxococcus xanthus is controlled by the espAB
RT   locus.";
RL   Mol. Microbiol. 34:714-725(1999).
```

Align seg 1/1 to: O35021 from: 1 to: 224

```
718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
    ||||||||||||||||||||||||||||||||
 58 AlaIleAsnProGlyAsnSerGlyGlyProLeu 68
```

seq_name: sp_bacteria:P72780

seq_documentation_block:
```
ID   P72780      PRELIMINARY;      PRT;   394 AA.
AC   P72780;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROTEASE HHOA.
GN   HHOA.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90900; BAA16795.1; -.
DR   MEROPS; S01.274; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 1.
DR   PRINTS; PR00834; PROTEASES2C.
KW   Protease.
SQ   SEQUENCE   394 AA;  41336 MW;  AD59D94811B8F57B CRC64;
```

alignment_scores:
```
        Quality:   11.00           Length:        11
        Ratio:      1.000          Gaps:           0
Percent Similarity: 100.000     Percent Identity: 100.000
```

alignment_block:
US-09-388-090-3 x P72780   ..

Align seg 1/1 to: P72780 from: 1 to: 394

```
718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
    ||||||||||||||||||||||||||||||||
231 AlaIleAsnProGlyAsnSerGlyGlyProLeu 241
```

seq_name: sp_bacteria:O06670

seq_documentation_block:
```
ID   O06670      PRELIMINARY;      PRT;   397 AA.
AC   O06670;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PUTATIVE SERINE PROTEASE.
GN   SPHTRA.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R801;
RA   Gasc A.A.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF000658; AAC45334.1; -.
DR   MEROPS; S01.273; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 1.
DR   PRINTS; PR00834; PROTEASES2C.
KW   Protease.
SQ   SEQUENCE   397 AA;  42287 MW;  755C13398ABBEE15 CRC64;
```

alignment_scores:
```
        Quality:   11.00           Length:        11
        Ratio:      1.000          Gaps:           0
Percent Similarity: 100.000     Percent Identity: 100.000
```

alignment_block:
US-09-388-090-3 x O06670   ..

Align seg 1/1 to: O06670 from: 1 to: 397

```
718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
    ||||||||||||||||||||||||||||||||
232 AlaIleAsnProGlyAsnSerGlyGlyProLeu 242
```

seq_name: sp_bacteria:O83752

seq_documentation_block:
```
ID   O83752      PRELIMINARY;      PRT;   398 AA.
AC   O83752;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PERIPLASMIC SERINE PROTEASE DO (HTRA-1).
GN   TP0773.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NICHOLS;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
DR   EMBL; AE001248; AAC65740.1; -.
DR   HSSP; P41140; 2SFA.
DR   MEROPS; S01.274; -.
DR   TIGR; TP0773; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001314; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00834; PROTEASES2C.
KW   Protease.
SQ   SEQUENCE   398 AA;  42186 MW;  8715EA05F454258D CRC64;
```

```
alignment_scores:
        Quality:   11.00              Length:    11
        Ratio:     1.000              Gaps:       0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x O83752      ..

Align seg 1/1  to: O83752  from: 1 to: 398

    718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
        |||||||||||||||||||||||||||||||||||
    223 AlaIleAsnProGlyAsnSerGlyGlyProLeu 233

seq_name: sp_plant:O04674

seq_documentation_block:
ID    O04674      PRELIMINARY;      PRT;    398 AA.
AC    O04674;
DT    01-JUL-1997 (TrEMBLrel. 04, Created)
DT    01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT    01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE    HTRA-LIKE PROTEIN.
GN    HTRA.
OS    Haematococcus pluvialis.
OC    Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC    Haematococcaceae; Haematococcus.
OX    NCBI_TaxID=44745;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=97324581; PubMed=9180686;
RA    Hershkovits G., Dubinsky Z., Katcoff D.J.;
RT    "A novel homologue of the prokaryotic htrA gene is differentially
RT    expressed in the alga Haematococcus pluvialis following stress.";
RL    Mol. Gen. Genet. 254:345-350(1997).
DR    EMBL; U58980; AAB61311.1; -.
DR    MEROPS; S01.279; -.
DR    MENDEL; 23986; Haepl;3069;23986.
DR    INTERPRO; IPR001254; -.
DR    INTERPRO; IPR001478; -.
DR    INTERPRO; IPR001940; -.
DR    PFAM; PF00089; trypsin; 1.
DR    PFAM; PF00595; PDZ; 1.
DR    PRINTS; PR00834; PROTEASES2C.
SQ    SEQUENCE   398 AA;   41469 MW;   D4D404273FB4309A CRC64;

alignment_scores:
        Quality:   11.00              Length:    11
        Ratio:     1.000              Gaps:       0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x O04674      ..

Align seg 1/1  to: O04674  from: 1 to: 398

    718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
        |||||||||||||||||||||||||||||||||||
    237 AlaIleAsnProGlyAsnSerGlyGlyProLeu 247

seq_name: sp_bacteria:P73940

seq_documentation_block:
ID    P73940      PRELIMINARY;      PRT;    416 AA.
AC    P73940;
DT    01-FEB-1997 (TrEMBLrel. 02, Created)
DT    01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT    01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE    PROTEASE HHOB.
GN    HHOB.
OS    Synechocystis sp. (strain PCC 6803).
OC    Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX    NCBI_TaxID=1148;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=97061201; PubMed=8905231;
RA    Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA    Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA    Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA    Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA    Tabata S.;
RT    "Sequence analysis of the genome of the unicellular cyanobacterium
RT    Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT    entire genome and assignment of potential protein-coding regions.";
RL    DNA Res. 3:109-136(1996).
DR    EMBL; D90911; BAA18006.1; -.
DR    MEROPS; S01.275; -.
DR    INTERPRO; IPR001254; -.
DR    INTERPRO; IPR001478; -.
DR    INTERPRO; IPR001940; -.
DR    PFAM; PF00089; trypsin; 1.
DR    PFAM; PF00595; PDZ; 1.
DR    PRINTS; PR00834; PROTEASES2C.
KW    Protease.
SQ    SEQUENCE   416 AA;   43196 MW;   DD32BB997CC6B25B CRC64;

alignment_scores:
        Quality:   11.00              Length:    11
        Ratio:     1.000              Gaps:       0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x P73940      ..

Align seg 1/1  to: P73940  from: 1 to: 416

    718 GCCATCAATCCGGGCAATTCCGGCGCGGCCCGCTG 750
        |||||||||||||||||||||||||||||||||||
    252 AlaIleAsnProGlyAsnSerGlyGlyProLeu 262

seq_name: sp_plant:Q9LK85

seq_documentation_block:
ID    Q9LK85      PRELIMINARY;      PRT;    439 AA.
AC    Q9LK85;
DT    01-OCT-2000 (TrEMBLrel. 15, Created)
DT    01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT    01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE    DEGP PROTEASE PRECURSOR.
OS    Arabidopsis thaliana (Mouse-ear cress).
OC    Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC    Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC    Brassicales; Brassicaceae; Arabidopsis.
OX    NCBI_TaxID=3702;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=COLUMBIA;
RA    Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL    Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN    [2]
RP    SEQUENCE FROM N.A.
RC    STRAIN=COLUMBIA;
RA    Nakamura Y.;
RT    "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT    Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT    TAC and BAC clones.";
RL    DNA Res. 7:217-221(2000).
DR    EMBL; AP000371; BAB02539.1; -.
DR    EMBL; AP001302; BAB02539.1; JOINED.
KW    Protease.
SQ    SEQUENCE   439 AA;   46673 MW;   05EB437DCE71A251 CRC64;

alignment_scores:
```

```
          Quality:      11.00              Length:        11
            Ratio:      1.000                Gaps:         0
Percent Similarity:   100.000    Percent Identity:   100.000

alignment_block:
US-09-388-090-3 x Q9LK85      ..

Align seg 1/1  to: Q9LK85  from: 1  to: 439

    718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
        ||||||||||||||||||||||||||||||||
    276 AlaIleAsnProGlyAsnSerGlyGlyProLeu 286

seq_name: sp_bacteria:Q9RTK4

seq_documentation_block:
ID  Q9RTK4       PRELIMINARY;      PRT;      441 AA.
AC  Q9RTK4;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE  PERIPLASMIC SERINE PROTEASE DO, PUTATIVE.
GN  DR1756.
OS  Deinococcus radiodurans.
OC  Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX  NCBI_TaxID=1299;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=R1;
RX  MEDLINE=20036896; PubMed=10567266;
RA  White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA  Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA  Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA  Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA  Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA  Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA  Fraser C.M.;
RT  "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT  radiodurans R1.";
RL  Science 286:1571-1577(1999).
DR  EMBL; AE002017; AAF11312.1; -.
HSSP; P00778; 1TAL.
TIGR; DR1756; -.
DR  INTERPRO; IPR001254; -.
DR  INTERPRO; IPR001478; -.
DR  INTERPRO; IPR001940; -.
DR  PFAM; PF00089; trypsin; 1.
DR  PFAM; PF00595; PDZ; 1.
DR  PRINTS; PR00834; PROTEASES2C.
KW  protease.
SQ  SEQUENCE  441 AA;  45145 MW;  D9E475AD081E0190 CRC64;

alignment_scores:
          Quality:      11.00              Length:        11
            Ratio:      1.000                Gaps:         0
Percent Similarity:   100.000    Percent Identity:   100.000

alignment_block:
US-09-388-090-3 x Q9RTK4      ..

Align seg 1/1  to: Q9RTK4  from: 1  to: 441

    718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
        ||||||||||||||||||||||||||||||||
    256 AlaIleAsnProGlyAsnSerGlyGlyProLeu 266

seq_name: sp_plant:Q9LU10

seq_documentation_block:
ID  Q9LU10       PRELIMINARY;      PRT;      448 AA.
AC  Q9LU10;
DT  01-OCT-2000 (TrEMBLrel. 15, Created)
DT  01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT  01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE  GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K13H13.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC  Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC  Brassicales; Brassicaceae; Arabidopsis.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=COLUMBIA;
RA  Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA  Tabata S.;
RT  "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL  Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR  EMBL; AB024023; BAA98101.1; -.
SQ  SEQUENCE  448 AA;  47492 MW;  A986FC1387670AFF CRC64;

alignment_scores:
          Quality:      11.00              Length:        11
            Ratio:      1.000                Gaps:         0
Percent Similarity:   100.000    Percent Identity:   100.000

alignment_block:
US-09-388-090-3 x Q9LU10      ..

Align seg 1/1  to: Q9LU10  from: 1  to: 448

    718 GCCATCAATCCGGCAATTCCGGCGGCCCGCTG 750
        ||||||||||||||||||||||||||||||||
    286 AlaIleAsnProGlyAsnSerGlyGlyProLeu 296

seq_name: sp_bacteria:O34358

seq_documentation_block:
ID  O34358       PRELIMINARY;      PRT;      449 AA.
AC  O34358;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT  01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE  YKDA.
GN  YKDA OR HTRA.
OS  Bacillus subtilis.
OC  Bacteria; Firmicutes; Bacillus/Clostridium group;
OC  Bacillus/Staphylococcus group; Bacillus.
OX  NCBI_TaxID=1423;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=168;
RA  Devine K.M.;
RL  Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=168;
RX  MEDLINE=98044033; PubMed=9384377;
RA  Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA  Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA  Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA  Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA  Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA  Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA  Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA  Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA  Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA  Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA  Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA  Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA  Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA  Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA  Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA  Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA  Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
```

```
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 1.
KW   Protease.
SQ   SEQUENCE   452 AA;   47656 MW;  AED4AFB1AC472361 CRC64;

alignment_scores:
         Quality:     11.00          Length:     11
         Ratio:       1.000          Gaps:        0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x P73354

Align seg 1/1  to: P73354   from: 1  to: 452

718 GCCATCAATCCGGGCAATTCCGGCGGCCCCGCTG 750
    ||||||||||||||||||||||||||||||||||
290 AlaIleAsnProGlyAsnSerGlyGlyProLeu 300
```

```
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ002571; CAA05570.1; -.
DR   EMBL; Z99110; CAB13147.1; -.
DR   MEROPS; S01.273; -.
DR   INTERPRO; IPR001254; -.
DR   INTERPRO; IPR001478; -.
DR   INTERPRO; IPR001940; -.
DR   PFAM; PF00089; trypsin; 1.
DR   PFAM; PF00595; PDZ; 1.
DR   PRINTS; PR00834; PROTEASES2C.
KW   Protease.
SQ   SEQUENCE   449 AA;   47713 MW;  E12B07A9018EE414 CRC64;

alignment_scores:
         Quality:     11.00          Length:     11
         Ratio:       1.000          Gaps:        0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-388-090-3 x O34358

Align seg 1/1  to: O34358  from: 1  to: 449

718 GCCATCAATCCGGGCAATTCCGGCGGCCCCGCTG 750
    ||||||||||||||||||||||||||||||||||
284 AlaIleAsnProGlyAsnSerGlyGlyProLeu 294

seq_name: sp_bacteria:P73354

seq_documentation_block:
ID   P73354      PRELIMINARY;    PRT;  452 AA.
AC   P73354;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   SERINE PROTEASE HTRA.
GN   HTRA.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90905; BAA17385.1; -.
DR   MEROPS; S01.273; -.
```

us-09-388-090-3.oligo7.rspt

Fri Mar 23 07:26:36 2001